Stock Market Reactions to the COVID-19 Pandemic

Steven J. Davis

Academic Keynote Talk, ABFER 8th Annual Conference

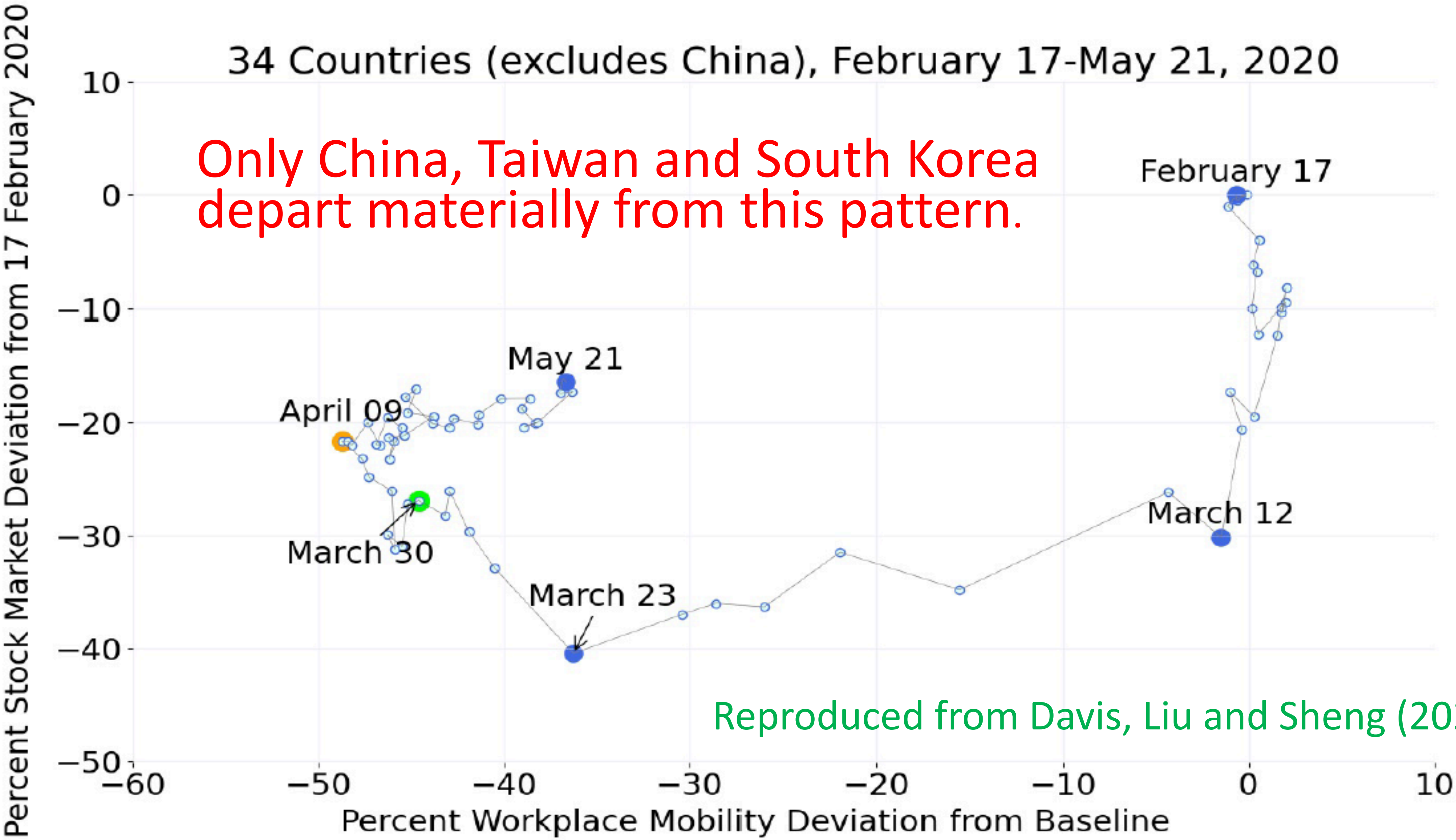
25 May 2021

This talk draws on research with Scott Baker, Nick Bloom, Stephen Hansen, Kyle Kost, Dingqian Liu, Marco Sammon, Cristhian Seminario-Amez, Xuguang Simon Sheng, and Tasaneeya Viratyosin.

Some Context

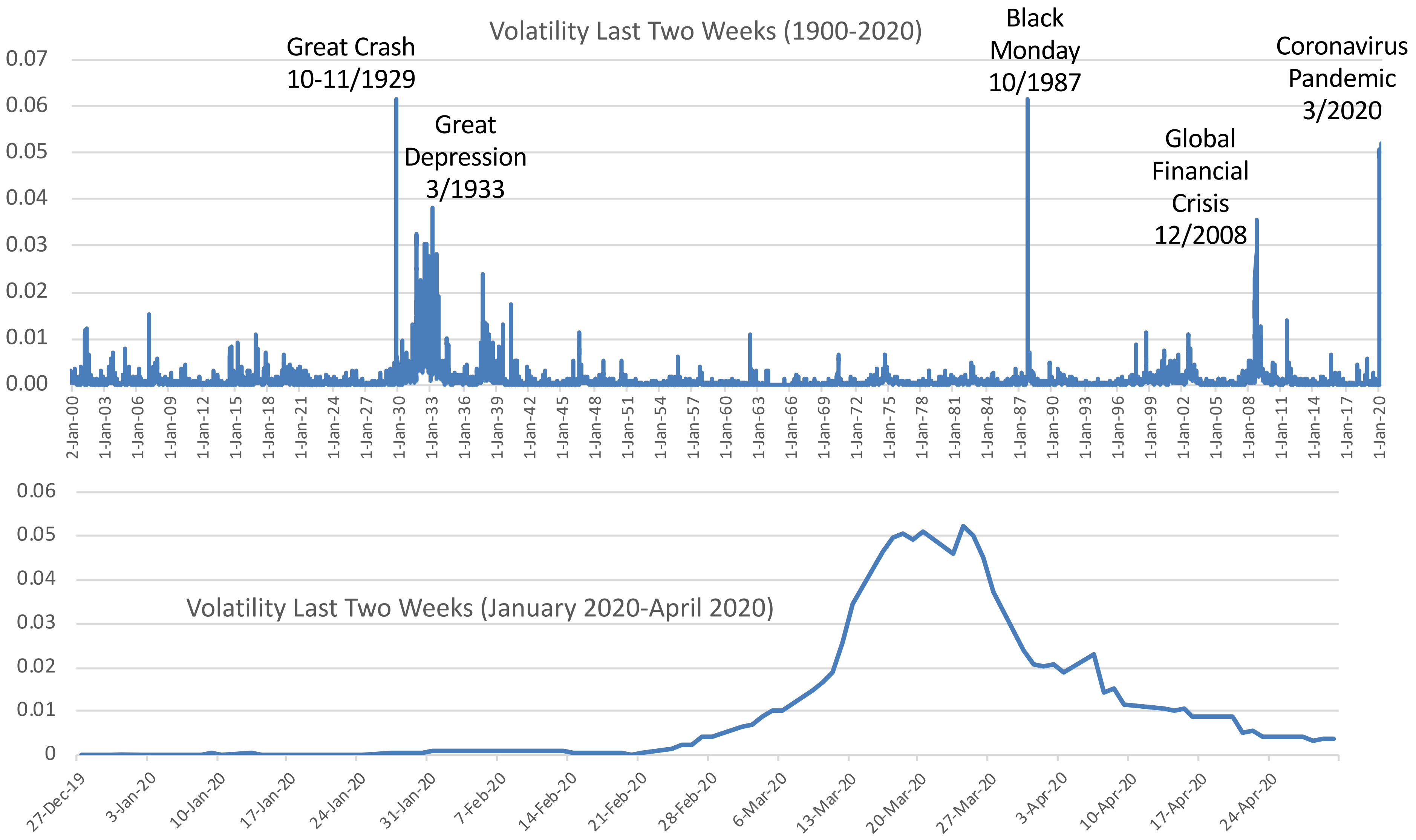
1. Tremendous stock market reactions to COVID-19:
   – Globally, stocks fell 40% from 17 Feb. to 23 March 2020.
   – Too big to rationalize with a standard asset-pricing model.
   – One of the great volatility episodes in the past 120 years.
   – More daily U.S. stock market moves > |2.5%| in March 2020 than ***any month since 1900***.
2. No previous pandemic, including the Spanish Flu, had remotely similar stock market effects in the United States or China.
3. Huge dispersion in firm-level equity return reactions to market-moving news in Feb-March 2020.
   – Re COVID-19 as a reallocation shock, see Barrero, Bloom and Davis (2020, 2021ab) and Schmidt and Papanikalaou (2021).

Time Path of Stock Prices and Workplace Mobility, Market-Cap Weighted Global Averages



Reproduced from Davis, Liu and Sheng (2021).

Realized U.S. Stock Market Volatility, January 1900 to April 2020



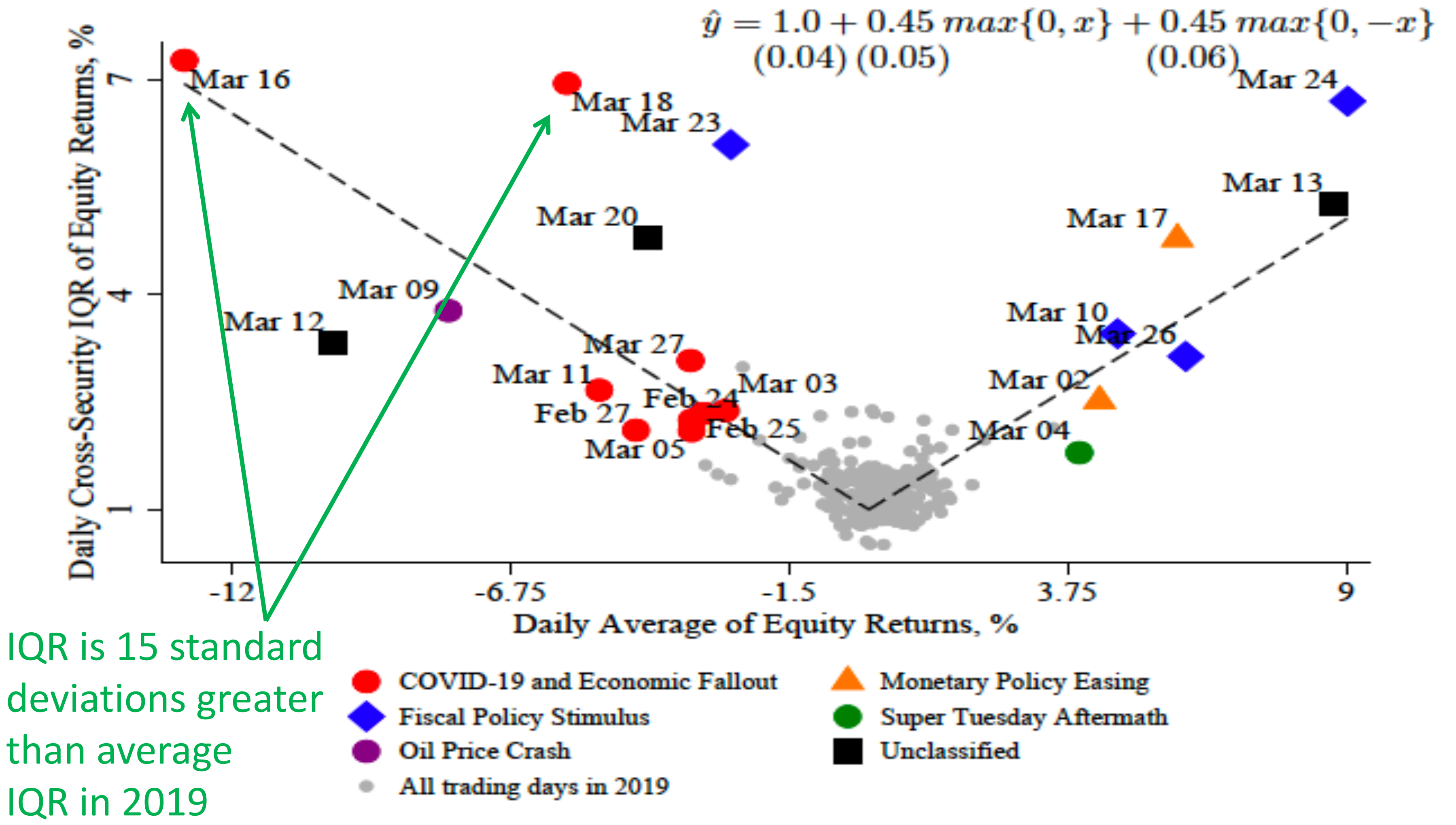
Notes: The sample period runs from 1/2/1900 to 4/30/2020. From December 1925 onwards, returns are computed using Yahoo Finance's 'adjusted close' series for the S&P 500 (^GSPC). Before that, returns are from the Global Financial Data extension of the Dow Jones Index. In both panels, we calculate realized volatility as the sum of squared returns over the past 10 trading days. Reproduced from Baker, Bloom, Davis, Kost, Sammon and Virayosin (2020).

The Unprecedented Stock Market Impact of the Coronavirus

	Number of Daily U.S. Stock Market Jumps Greater than |2.5%|	Number Attributed to Economic Fallout of Pandemics	Number Attributed to Policy Responses to Pandemics
2 January 1900 to 21 February 2020	1,116	0	0
24 February 2020 to 30 April 2020	27	13.4	10.4

Note: Tabulated from results in Baker, Bloom, Davis and Sammon (2020), who consider all daily jumps in the U.S. stock market greater than 2.5%, up or down, since 1900. They classify the reason for each jump into 17 categories based on human readings of next-day (or same-evening) accounts in the *Wall Street Journal* (and *New York Times* in 2020). Fractional counts arise when newspapers differ in their jump attribution or human readers differ in their classification of the attribution. Number Attributed to Economic Fallout of Pandemics includes jumps on 3/12 and 3/16 that a subset of coders classified as Macroeconomic Outlook. It's clear from reading these articles that the journalist regarded the deterioration in the Macroeconomic Outlook as due to the spread of the coronavirus.

Tremendous Dispersion in U.S. Firm-Level Stock Price Reactions to COVID News



Reproduced From Davis, Hansen and Seminario-Amez (2021)

Classifications of jump days from Baker et al. (2021)

Value-Weighted Mean and Cross-Sectional IQR of U.S. Equity Returns, Daily for 2019 and Large Daily Jumps in February and March 2020.

Davis, Hansen & Seminario-Amez (2021)

- Characterize firm-level shock exposures using the "Risk Factors" texts in annual 10-K filings.
- Use exposure measures to explain firm-level return reactions to market-moving news (daily market-level jumps > |2.5%|)
  - Focus on jumps from late February to end of April (2020).
- Implement and compare two text-analytic approaches:
  - Expert-curated dictionaries (e.g., Tetlock (2007), Loughran-McDonald (2011), Baker-Bloom-Davis (2016))
  - Taddy's (2013) Multinomial Inverse Regression (MNIR), a form of supervised machine learning (ML)
- Develop a hybrid approach to uncover new exposures and sharpen explanations/interpretations of firm-level returns.

Using the Text in Regulatory Filings to Quantify Firm-Level Exposures and Characterize Return Drivers

To explain firm-level returns, we use the *Risk Factors* discussion of 10-K regulatory filings.

These texts discuss factors that generate uncertainty in future earnings; exhaustive due to their legal status.

RF corpus for 2,133 companies for the 2010-2016 time period.

Key idea: RF content that explains abnormal returns on jump dates reveals channels through which future earnings react to macro shock.

Dictionary Approach

- Dictionaries from Baker et al. (2019), who expand on ones developed by BBD (2016) and Davis (2017).
  - 16 dictionaries cover aspects of economic and financial conditions
  - 20 pertain to policy areas.
  - Each dictionary contains terms that effectively define the category.
- In aggregate, the dictionaries contain 244 distinct terms that appear 1.4 million times (2.4% of the entire *RF* corpus).
- The *RF* texts for a given firm contain 28 distinct dictionary terms on average (standard deviation of 10) and 642 instances of dictionary terms (standard deviation of 620).
- To quantify a firm's exposure to a given category, we identify sentences in its *RF* texts that contain at least one term in the corresponding dictionary.
- After computing the fraction of such sentences in each of the firm's *RF* texts, we average over years for the firm. This yields 36 firm-level exposure values, one for each category.

Dictionary Examples

Inflation: {cpi, inflation, gold, silver}

Commodity Markets: {wheat, corn, sugar, cotton, beef, pork, petroleum, oil, coal, natural gas, biofuel, ethanol, steel, copper, zinc, tin, platinum, gold, metal, silver, aluminum, lead, commodity exchange, nymex, mercantile exchange, gas pipeline}

Monetary Policy: {monetary policy, money supply, open market operations, discount window, quantitative easing, central bank, federal reserve, the fed, european central bank}

Taxes: {taxes, tax, taxation, taxed, vat, accelerated depreciation, fiscal cliff, internal revenue service}

Trade Policy: {tariff, dumping, world trade organization, north american free trade agreement, international trade commission}

Firm-Level Return Regressions (Least Squares)

$$\mathrm{Abn}_{it} = \sum_{j=1}^{J} \beta_j \, \mathrm{RExp}_i^j + \beta_{J+1} \, \mathrm{Leverage}_i + \beta_{J+2} \, \log(\mathrm{Mcap}_{it}) + \gamma_{s(i)} + \epsilon_{it},$$

Where:

- $\mathrm{Abn}_{i,t}$ is the abnormal return of firm *i* on jump-day *t,* or its average abnormal return on a collection of jump dates.
- RExp_i^j is firm *i'*s exposure to category *j* risks/shocks.
- Leverage and Mcap (market capitalization) are controls.
- $\gamma_{s(i)}$ is a vector of industry effects (NAICS2 or NAICS4 level).

Under the dictionary approach, $\beta_j \, \mathrm{RExp}_i^j$ captures the effect of firm *i'*s exposure to category *j* on its one-day abnormal equity return. Fit to data for "jump" days in February and March 2020.

Jump Classification ⟶ Dependent Variable: Abn_{it}	(1) COVID-19 and Its Fallout		(2) Monetary Policy Easing	
General Economic Categories				
Inflation	−0·21	(−2·5)	0·92	(4·7)
Interest Rates			0·78	(5·4)
Credit Indicators	−0·29	(−4·1)	−0·68	(−3·4)
Labor Markets				
Real Estate Markets			0·51	(2·2)
Business Investment and Sentiment				
Consumer Spending and Sentiment			−0·36	(−1·9)
Commodity Markets			−0·41	(−2·0)
Healthcare Matters			0·62	(2·0)
Litigation Matters				
Competition Matters			−0·37	(−1·8)
Intellectual Property Matters	0·45	(6·2)		
Policy-Related Categories				
Taxes	−0·28	(−2·1)		
Entitlement and Welfare Programs	−0·49	(−2·9)		
Monetary Policy				
Financial Regulation	0·12	(1·8)		
Competition Policy			0·32	(2·0)
Intellectual Property Policy				
Energy and Environmental Regulation	−0·19	(−2·2)	−0·31	(−2·1)
Housing and Land Management				
Other Regulation			0·25	(3·2)
Healthcare Policy	0·31	(1·9)		
Transportation, Infrastructure, Utilities	−0·16	(−2·6)		
Elections and Political Governance				
Financial Controls				
Log Market Cap	0·53	(7·3)	0·73	(3·0)
Leverage	−0·42	(−3·0)	−0·85	(−2·8)
Observations [Adjusted R^2]	2155	[0.329]	2155	[0.232]

This table excerpt reports our fitted dictionary-based regression models for the ***9 pandemic fallout days*** and the ***2 monetary policy easing days***

Multinomial Inverse Regression Approach, 1

- MNIR treats the *RF* texts for each firm as a *bag-of-words* represented by a *V*-dimensional vector x_i of terms or ``features.''
- $x_{i,v}$ is the count of term *v* for firm *i*, and *V* = 18,911 is the number of unique terms in our *RF* corpus.
- At the firm level, the average number of nonzero elements in x_i is 2,245, with a standard deviation of 891.

Multinomial Inverse Regression Approach, 2

MNIR posits $\mathbf{x}_i \sim \mathrm{MN}(\mathbf{q}_i, N_i)$, where $\mathbf{q}_i$ is a multinomial V-dimensional probability vector and N_i is the total number of terms in firm i's RF texts (i.e., $N_i = \sum_v x_{i,v}$). The probability of feature v for firm i is

$$q_{i,v} = \frac{\exp(a_v + \mathbf{y}_i^T \mathbf{b}_v)}{\sum_v \exp(a_v + \mathbf{y}_i^T \mathbf{b}_v)}, \tag{2}$$

where $\mathbf{y}_i = (\mathrm{Abn}_i, \mathbf{c}_i)$ contains firm-i abnormal returns on a given day or collection of days and firm controls $\mathbf{c}_i \in \mathbb{R}^P$. (We suppress time subscripts here.) a_v is a parameter that controls for the baseline frequency of term v in the corpus, and $\mathbf{b}_v$ is a $P+1$ vector of coefficients that describe how firm observables map to the probability that term v appears in the RF texts.

Multinomial Inverse Regression Approach, 3

Equation (2) describes a multinomial logistic regression over *V* categories, which we fit to 2,155 observations per jump day, one per firm. Here, we model the probability that a particular term in *V* appears in a random draw from the firm's *RF* texts.

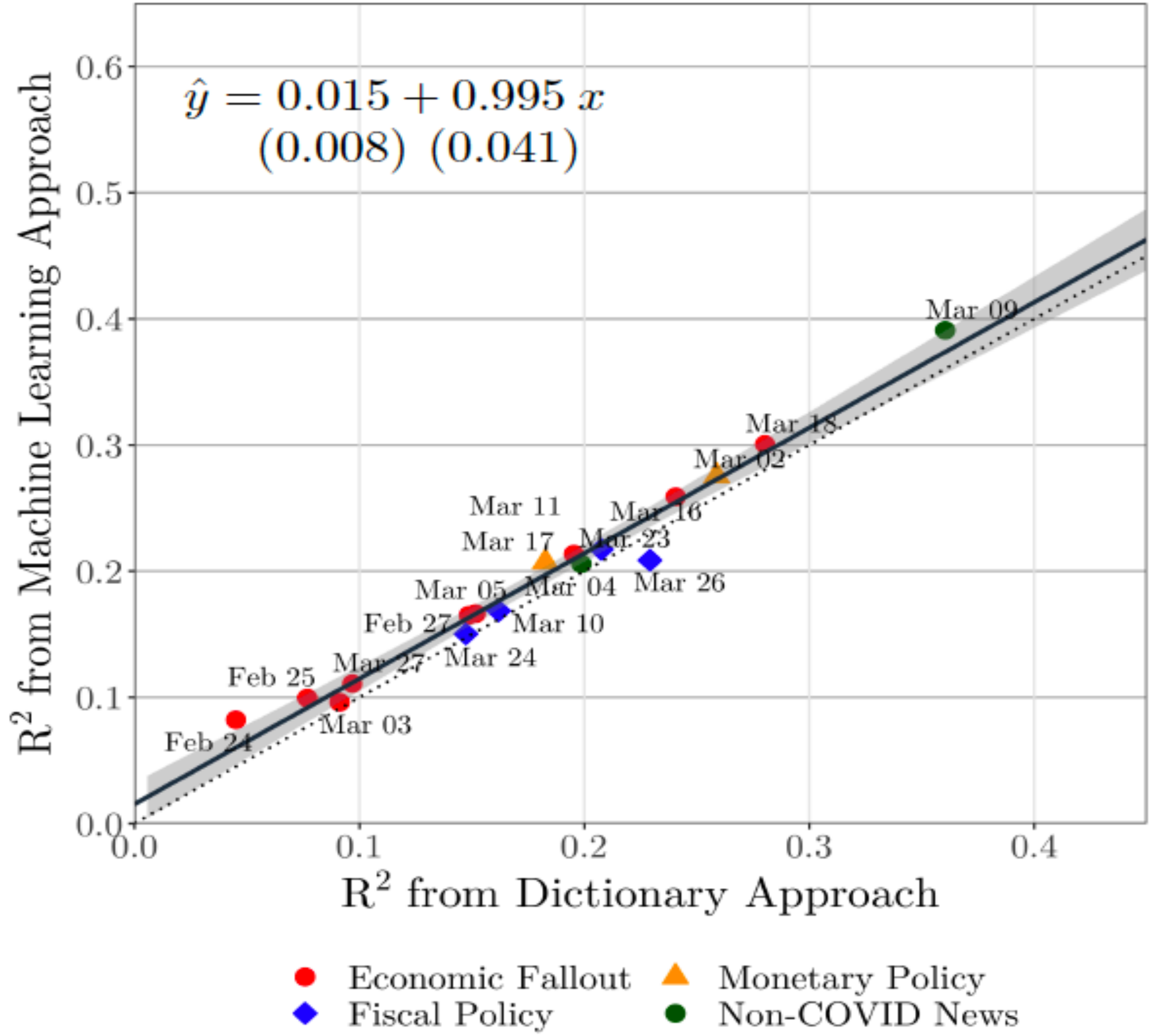
We fit (2) using Bayesian regularization methods with a Gamma-Laplace prior structure on the regression coefficients. The prior trades off goodness-of-fit and model complexity, maximizing an information criterion to avoid over-fitting. See Taddy (2013, 2015) for details.

Multinomial Inverse Regression Approach, 4

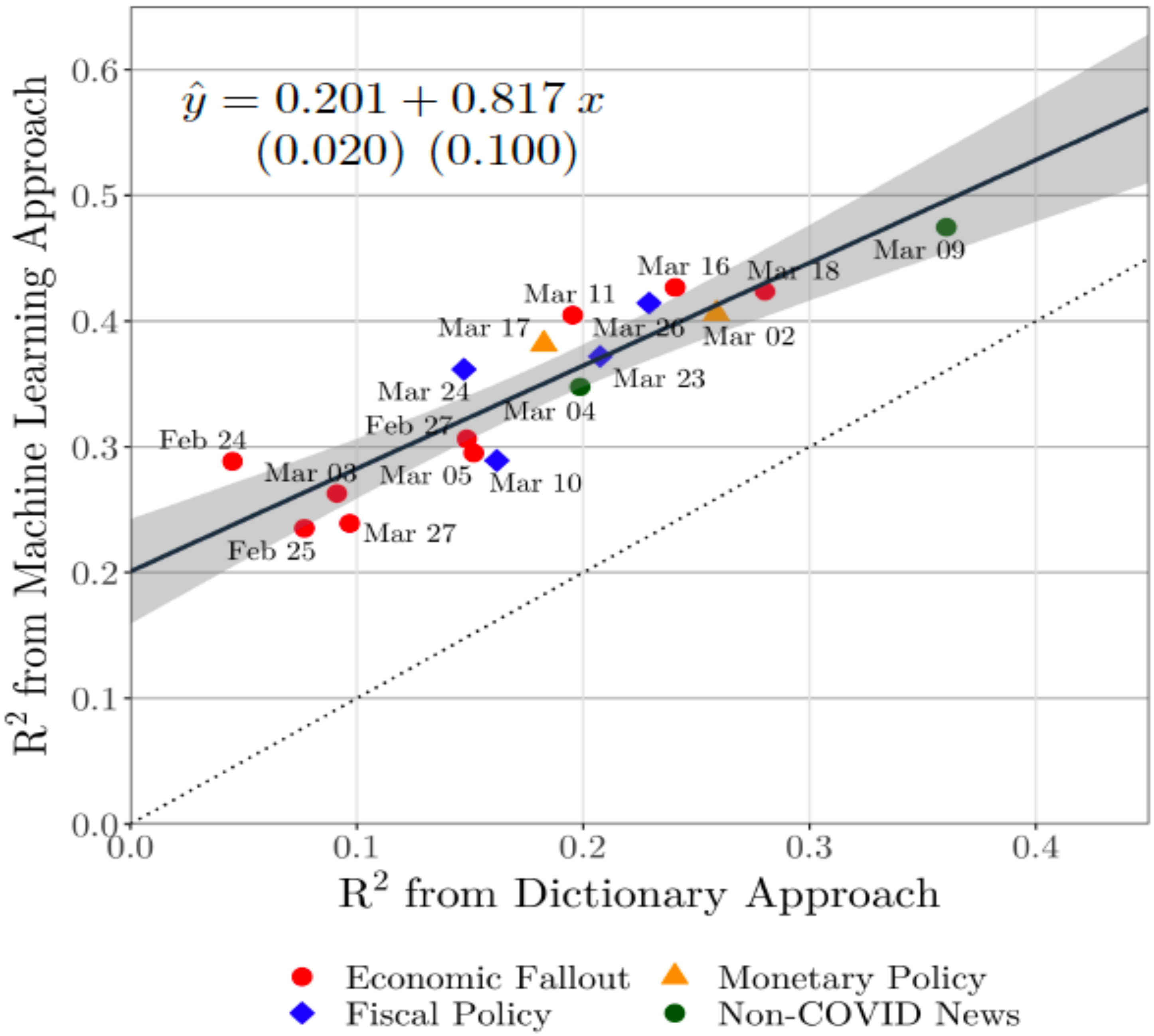
To move from (2) to a ***forward regression*** with Abn_i as the dependent variable, we follow Taddy (2013) and

define a *sufficient reduction projection* $z_i = \sum_v x_{i,v} b_{1,v}$ with the property $Abn_i \perp \mathbf{x}_i \mid z_i, N_i, \mathbf{c}_i$. Thus, conditional on the scalar projection z_i, the high-dimensional raw data contain no extra predictive information for returns. This result does not specify the functional form for relating z_i to Abn_i in a forward regression, but it says we can model Abn_i as a function of z_i, N_i, $\mathbf{c}_i$, while disregarding $\mathbf{x}_i$.

How Much Fit Gain from MNIR, and Why?

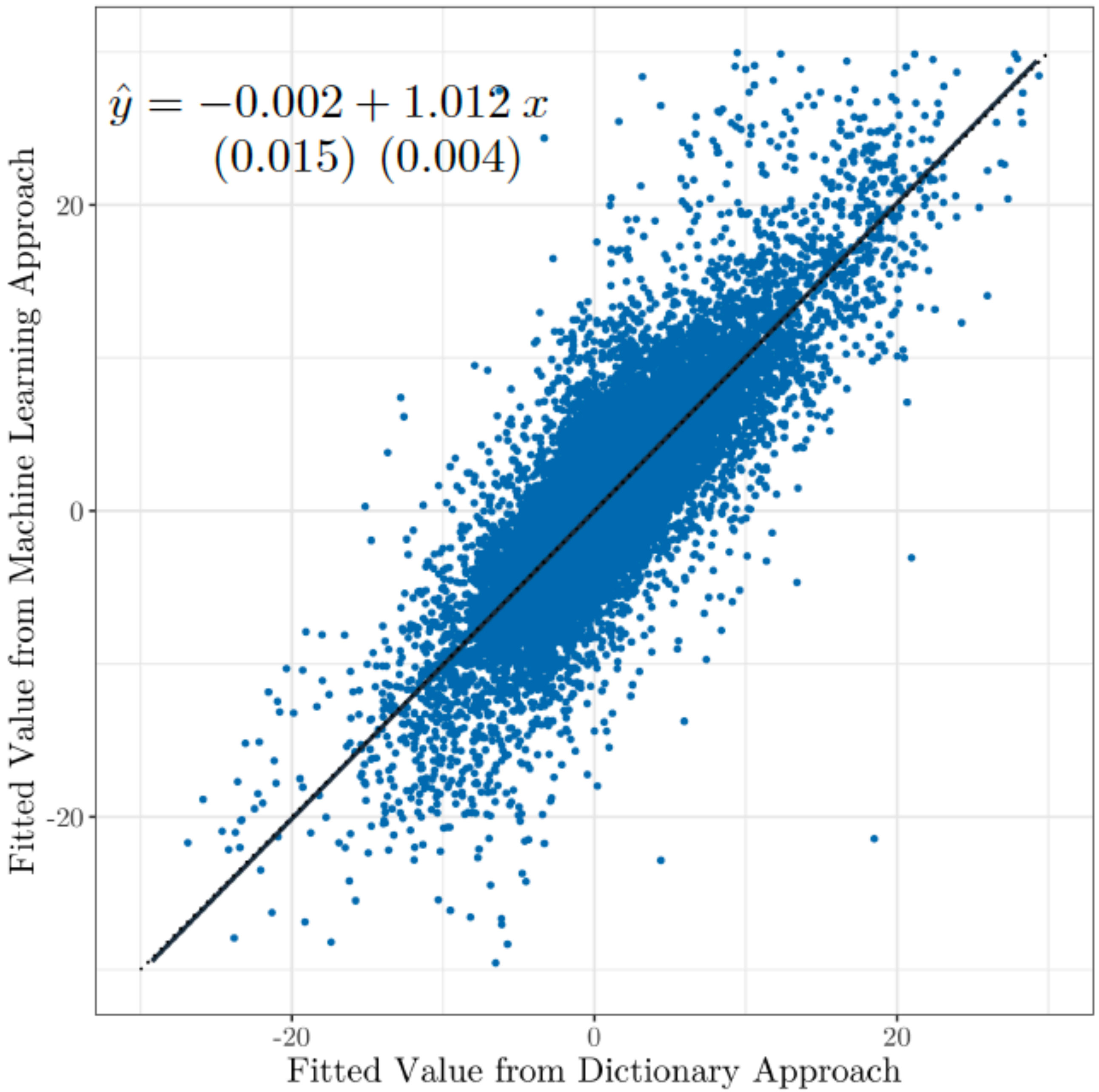


(a) Restricted Feature Space in MNIR



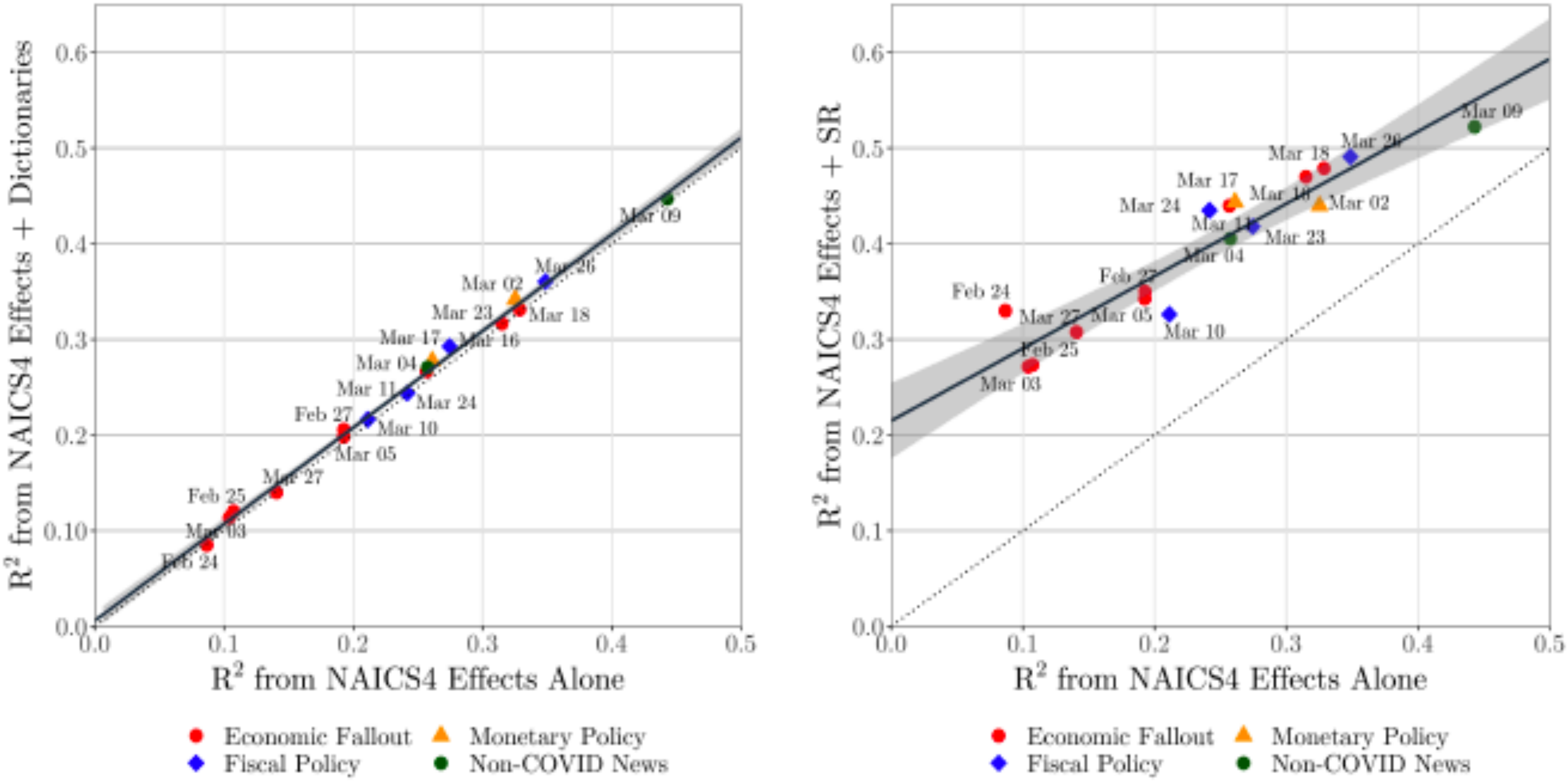
(b) Full Feature Space in MNIR

Comparing Firm-Level Predictions



MNIR Captures Much More than Standard Industry Codes

Dependent variable in regressions: 1-day firm-level abnormal Returns. All regressions include firm-level financial controls and NAICS4 fixed Effects.



(a) Dictionary Method (b) Machine Learning Method

Figure 5: Improvement in R^2 beyond Narrow Industry Codes

What Accounts for the Fit?

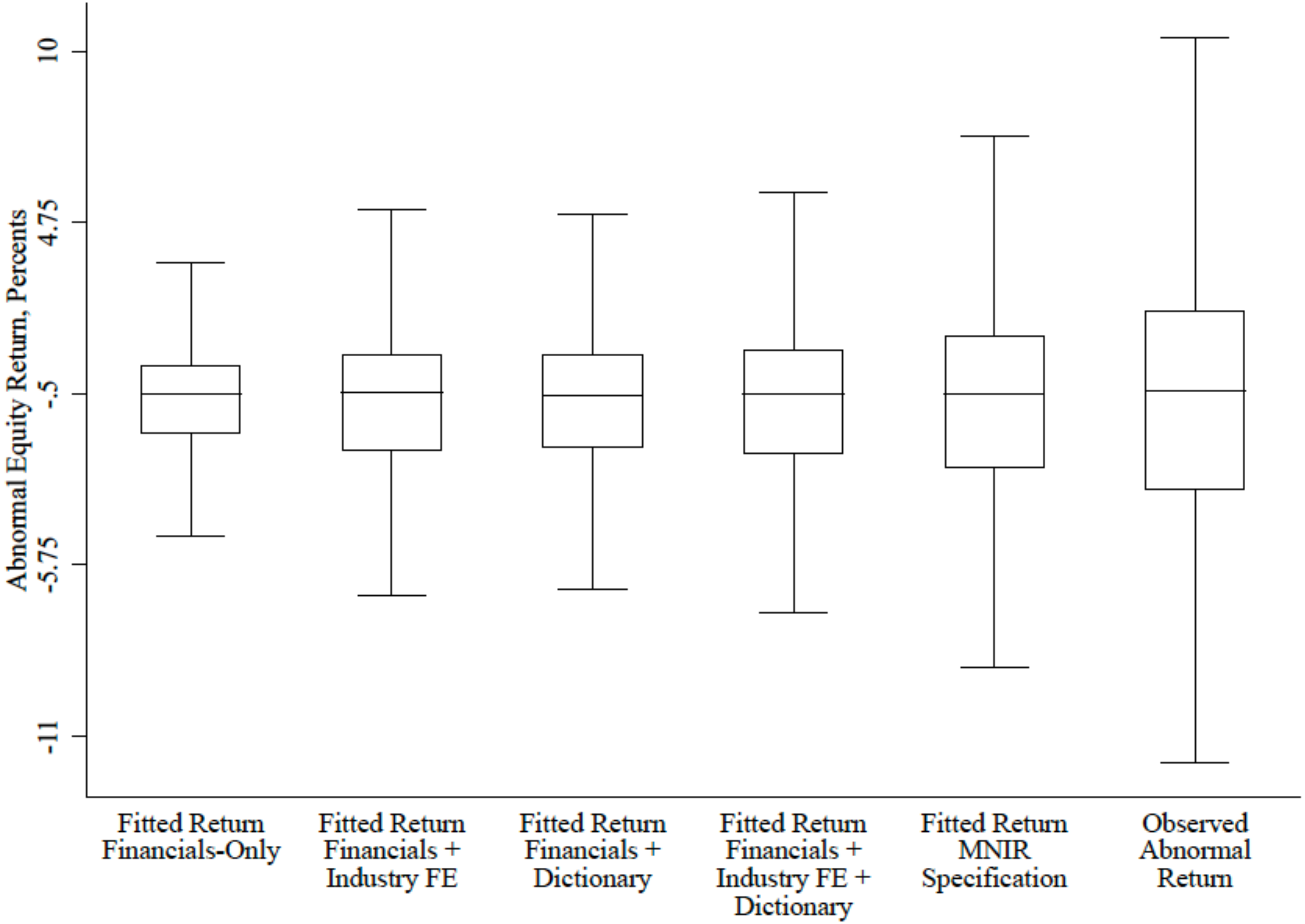


Figure 4: Range of Firm-Level (Fitted) Returns

Why A Hybrid Approach?

- MNIR: impressive predictive performance, but raw results help little to explain/understand what drives the structure of firm-level returns.
- By combining MNIR with a semi-automated process for constructing new exposure categories, we:
  - Sidestep laborious task of constructing expert-curated dictionaries (and lean much less on domain expertise).
  - Obtain more granular exposure categories & measures.
  - Fit firm-level returns better than models generated by expert-curated dictionaries.
  - Obtain straightforward regression results that have much greater interpretive value than MNIR alone.

How Our Hybrid Approach Works

Step 1: Select seed terms

To construct risk exposures, we start with "seeds" drawn from (a) terms with large MNIR coefficients, $|\hat{b}_{1,v}|$; and (b) terms with large tf-idf weighted MNIR coefficients, $|\hat{b}_{1,v}| x_v \log(\frac{2133}{df_v})$, where x_v is the count of term v in the *RF* corpus. We work with 45 seeds that reflect both positive and negative return reactions and that appear to cover the main exposures surfaced by our MNIR model fit to pandemic-related jump days.

Step 2: Build term sets corresponding to each seed

Associate each seed with terms in the *RF* corpus that meet two criteria:

1. High specificity and same sign: Among terms v with an MNIR coefficient of the same sign as the one for the seed, we select those with $|\hat{b}_{1,v}|x_v \log(\frac{2133}{df_v}) > 200$.
2. High contextual similarity, as measured by cosine similarity of the embedding vectors: In practice, we require a term's embedding vector to have a cosine similarity greater than 0.4 with that of the seed.

Step 3: Manual pruning of term sets and, in a few cases, combining term sets

Exposure Categories for Pandemic Fallout Dates: Hybrid Approach

Seed	Name	Retained Terms	Dropped Terms
preclinical	Drug Trials	43	0
ecommerce	Ecommerce	12	11
optics	Electronic Components and Devices	74	20
wheat	Foodstuffs	27	2
china	Foreign Countries	62	0
medicare	Health Insurance	35	0
investment funds	Investment Funds	15	0
manufacturing	Manufacturing	35	5
steel	Metal Products	21	0
coal	Power Generation	13	0
tantalum	Raw Metals and Minerals	11	3
semiconductor	Semiconductors	15	5
games	Video Games	21	4
cloud	Web-Based Services	23	2
bank	Banking	40	0
fdic	Deposits	20	0
vessels	Shipping Containers	12	1
freight	Transportation	21	0
solutions	Software Services	65	8
software	Software and Hardware Products	56	9

(b) Positive Exposures

Seed	Name	Retained Terms	Dropped Terms
advertisers	Advertizing	9	5
biodiesel	Alternative Energy	10	17
card	Card Payments	25	0
clearing house	Clearing Houses	3	0
hotels	Commercial Property	18	3
display	Display Technology	16	13
unrealized loss position	Financial Management	15	0
yen	Foreign Exchange	5	0
franchisees	Franchising	13	0
gaming	Gambling	5	0
gold	Gold and Silver	2	5
surgeons	Healthcare Providers	6	1
reinsurance	Insurance	23	0
mortgage	Mortgages	44	0
reit	REITs	29	0
homebuilding	Residential Construction	4	0
restaurants	Restaurants	3	13
retail	Traditional Retail	26	9
workforces	Workforce	2	0
aircraft	Aircraft and Airlines	10	10
travel	Travel	11	6
satellite	Communications	22	0
newspapers	Traditional Media	20	3
pipelines	Energy Infrastructure	26	11
oil	Oil and Gas	11	0

(a) Negative Exposures

Step 4: Compute firm-level exposures

Lastly, we compute firm-level exposures to each category j and its associated term set, $L(j)$, as $z_i^j = \sum_{v \in L(j)} x_{i,v}|b_{1,v}|$, which captures the part of the sufficient reduction projection that derives from terms in $L(j)$. Table B.3 reports descriptive statistics for

Step 5: Run the same type of regressions as when using expert-curated dictionaries to generate firm-level exposures.

$$\mathrm{Abn}_{it} = \sum_{j=1}^{J} \beta_j \, \mathrm{RExp}_i^j + \beta_{J+1} \, \mathrm{Leverage}_i + \beta_{J+2} \log(\mathrm{Mcap}_{it}) + \gamma_{s(i)} + \epsilon_{it},$$

How Well Do Our Return Regressions Fit?

Empirical Approach	Adjusted R-Squared Value in Firm-Level Abnormal Return Regression	
	Nine Pandemic Fallout Days	Day after Super Tuesday Election
Use expert-curated dictionaries and *RF* texts to quantify firm-level exposures.	0.33	0.20
Supervised Machine Learning: Forward regression in MNIR approach of Taddy.	0.50	0.35
Hybrid Approach: Use MNIR to build limited number of new dictionaries that yield firm-level risk exposure measures	0.41	0.24

Dependent Variable: Abn_{it}	(1) NAICS-2 Fixed Effects		(2) NAICS-2 Fixed Effects		(3) NAICS-4 Fixed Effects	
Exposures						
Advertizing	−0·09	(−2·4)	−0·10	(−2·2)	−0·12	(−3·0)
Alternative Energy	−0·10	(−6·8)	−0·09	(−8·7)	−0·05	(−1·9)
Card Payments	−0·14	(−3·3)	−0·12	(−3·2)	−0·17	(−4·8)
Clearing Houses	−0·10	(−9·7)				
Commercial Property					−0·15	(−2·3)
Financial Management	−0·23	(−11·5)	−0·24	(−12·8)	−0·29	(−3·5)
Foreign Exchange	−0·07	(−3·9)	−0·06	(−4·0)	−0·05	(−2·7)
Franchising	−0·10	(−1·8)	−0·12	(−3·2)	−0·15	(−2·2)
Gambling	−0·23	(−2·6)	−0·23	(−2·7)	−0·33	(−4·6)
Gold and Silver	−0·28	(−16·8)	−0·28	(−22·1)	−0·32	(−11·4)
Healthcare Providers	−0·14	(−6·5)	−0·12	(−7·8)		
Insurance	0·04	(2·1)	0·05	(2·4)		
Mortgages	−0·11	(−3·3)	−0·13	(−5·6)		
REITs	−0·39	(−4·8)	−0·39	(−4·5)		
Residential Construction	−0·37	(−14·0)	−0·33	(−12·0)	−0·22	(−2·5)
Restaurants	−0·22	(−4·6)	−0·25	(−4·4)	−0·21	(−3·2)
Traditional Retail	−0·33	(−6·3)	−0·37	(−7·2)	−0·28	(−3·6)
Workforce	−0·19	(−3·1)	−0·20	(−2·9)	−0·20	(−3·3)
Aircraft + Travel	−0·24	(−2·7)	−0·25	(−2·9)		
Communications + Trad Media	−0·09	(−2·4)	−0·09	(−2·3)	−0·11	(−2·9)
Energy Infr + Oil and Gas	−0·31	(−5·1)	−0·28	(−4·8)	−0·19	(−3·9)
Drug Trials	0·16	(11·4)	0·15	(10·7)	−0·04	(−2·7)
Ecommerce	0·15	(3·0)	0·15	(3·4)	0·14	(2·6)
Electronic Components and Devices	0·09	(4·1)	0·11	(4·2)	0·14	(3·6)
Foodstuffs	0·17	(4·3)	0·15	(4·9)	0·15	(4·8)
Foreign Countries	0·23	(2·7)	0·16	(1·8)		
Investment Funds	0·22	(14·8)	0·22	(16·5)	0·21	(13·0)
Metal Products					−0·08	(−1·7)
Raw Metals and Minerals	0·29	(7·9)	0·28	(10·3)	0·26	(4·7)
Semiconductors					−0·07	(−2·0)
Video Games	0·12	(4·1)	0·10	(12·3)	0·11	(8·8)
Web-Based Services	0·22	(3·8)	0·20	(3·4)	0·21	(3·9)
Banking + Deposits	0·18	(5·4)	0·19	(5·1)	0·18	(4·0)
Financial Controls						
Log Market Cap	0·46	(4·4)	0·44	(4·1)	0·50	(6·2)
Leverage	−0·34	(−3·0)	−0·26	(−2·6)	−0·14	(−1·4)
Observations [Adjusted R^2]	2155	[0.410]	1868	[0.433]	1868	[0.470]

This table shows results of the daily abnormal returns regression fit to the nine ***pandemic fallout jump days*** using the hybrid approach. Apologies for the small font.

Dependent Variable: Abn_{it}	(1) NAICS-2 Fixed Effects	
Exposures		
Aircraft		
Card Payments	−0·04	(−2·4)
Financial Instruments	−0·15	(−3·2)
Foodstuffs	−0·11	(−4·7)
Gambling	−0·20	(−7·4)
Hotels	−0·25	(−8·9)
Industrial Metals	−0·09	(−1·8)
Motor Vehicles		
Power Generation	−0·19	(−4·2)
Shipping	−0·21	(−4·8)
Traditional Media	−0·15	(−8·0)
Transportation	−0·08	(−3·9)
Asset Mngmt + Financial Mngmt	−0·19	(−9·5)
Banking + Financial Regul	−0·18	(−7·5)
Drilling Act + Fracking	−0·19	(−2·0)
Construction	0·22	(2·4)
Drugs	0·13	(3·0)
Electronic Communication	0·28	(3·7)
Foreign	0·08	(2·0)
Franchising	0·11	(2·6)
Government Contracting		
Insurance	0·13	(8·7)
Metals	0·16	(3·7)
Military	0·09	(2·6)
REITs	0·43	(8·7)
Rental Market	0·26	(3·1)
Utilities	0·18	(7·6)
Waste	0·16	(6·2)
Ecomm + Health Ins + Subsidies	0·20	(4·0)
Gov Healthcare + Healthcare Supp	0·30	(1·8)
Financial Controls		
Log Market Cap	0·63	(4·8)
Leverage	−0·10	(−0·8)
Observations [Adjusted R^2]	2155	[0.242]

This table shows results of the daily abnormal returns regression fit to the ***jump day following the Super Tuesday primary elections*** using the hybrid approach.

Summary of Results

1. Bad COVID-19 news lowers returns for firms with high exposures to travel, traditional retail, aircraft production and energy supply – directly and via downstream demand linkages – and raises them for firms with high exposures to healthcare policy, e-commerce, web services, drug trials and materials that feed into supply chains for semiconductors, cloud computing & telecom.
2. Monetary and fiscal policy responses strongly impact firm-level returns as well but differently than pandemic news.
3. Super Tuesday (a huge win for Biden) drove negative returns for firms with high exposure to hotels, gambling, fracking, and financial management; and positive returns for firms with high exposure to healthcare, REITs, property rentals, communications and construction.

Summary of Results, 2

4. Despite major methodological differences, dictionary approach and MNIR yield highly congruent predictions of firm-level returns.
5. By operating on a vastly larger feature space, MNIR outperforms with respect to goodness-of-fit.
    - Our dictionary-based model explains 1/3 of the (huge) firm-level abnormal return variation on pandemic jump days, while MNIR explains 1/2.
6. Our hybrid approach outperforms dictionary approach in terms of model fit and, unlike MNIR, yields readily interpretable results.
7. Our text-based models of firm-level abnormal returns have strong predictive content for future corporate earnings surprises.

Text-Based Models Fit to Feb-March 2020 Data Predict 2020 Q3 Earnings Surprises Relative to 2019 Q1 Forecasts

	Dependent variable:				
	Earnings Surprise				
	(1)	(2)	(3)	(4)	(5)
SRP (Pandemic Fallout)		0.320**			0.313***
		(0.134)			(0.094)
SRP (Monetary Policy)			0.325***		0.122**
			(0.051)		(0.056)
SRP (Super Tuesday)				0.323***	0.279***
				(0.065)	(0.082)
Leverage	0.0002	0.025	0.008	−0.013	0.016
	(0.069)	(0.066)	(0.068)	(0.068)	(0.064)
Log Market Cap	0.106***	0.085**	0.106***	0.131***	0.107***
	(0.036)	(0.036)	(0.031)	(0.036)	(0.033)
Observations	1,507	1,507	1,507	1,507	1,507
NAICS4 Effects	Y	Y	Y	Y	Y
R^2	0.211	0.229	0.244	0.242	0.269
Adjusted R^2	0.143	0.163	0.179	0.176	0.204
Note:					*p<0.1; **p<0.05; ***p<0.01

Concluding Remarks

The pandemic-induced return reactions in February-March 2020 that we uncover foretell shifts in the real economy. For example:

- Major job losses in the traditional retail sector, employment gains at online shopping and delivery firms, a persistent collapse in air travel, job cuts in aircraft production, numerous bankruptcies among oil and gas companies, a collapse of advertizing revenue in print media, and surging demand for cloud computing.
- Evidence that COVID-19 accelerated ongoing shifts to digital services and remote interactions across a host of activities. E.g.., the share of new U.S. patent applications that advance technologies to support video conferencing, telecommuting, remote interactivity, and working from home doubled in the wake of the pandemic.

Concluding Remarks, 2

Although often seen as methodological alternatives, our analysis suggests these expert-curated dictionary methods and supervised machine learning are complements as much as substitutes.

By combining elements of both, we obtain rich models that (a) fit better than models based on expert-curated dictionaries, (b) uncover new, empirically relevant exposure categories missed by the curated dictionaries and, at the same time, (c) deliver interpretable patterns in the estimated structure of firm-level returns.

This last feature pushes the supervised ML approach from prediction to interpretation.

Extra Slides

A Standard Asset-Pricing Model

Barro (2006) posits an endowment economy with a representative agent who has time-separable, isoelastic preferences over consumption. Log output evolves exogenously as a random walk with drift:

$$ln(A_{t+1}) \;=\; ln(A_t) \;+\; \gamma + u_{t+1} \;+ v_{t+1} \qquad (3)$$

where the drift $\gamma \geq\; 0$, u_{t+1} is i.i.d. normal with mean 0 and variance σ^2, and v_{t+1} picks up low-probability disaster shocks. Barro shows that the price of a one-period equity claim at *t* is

$$P_{t1} = A_t e^{-\rho-(\theta-1)\gamma+(1/2)(\theta-1)^2\sigma^2}[e^{-p} + (1 - e^{-p})\times\mathrm{E}\{(1 - b)^{1-\theta}\}] \qquad (4)$$

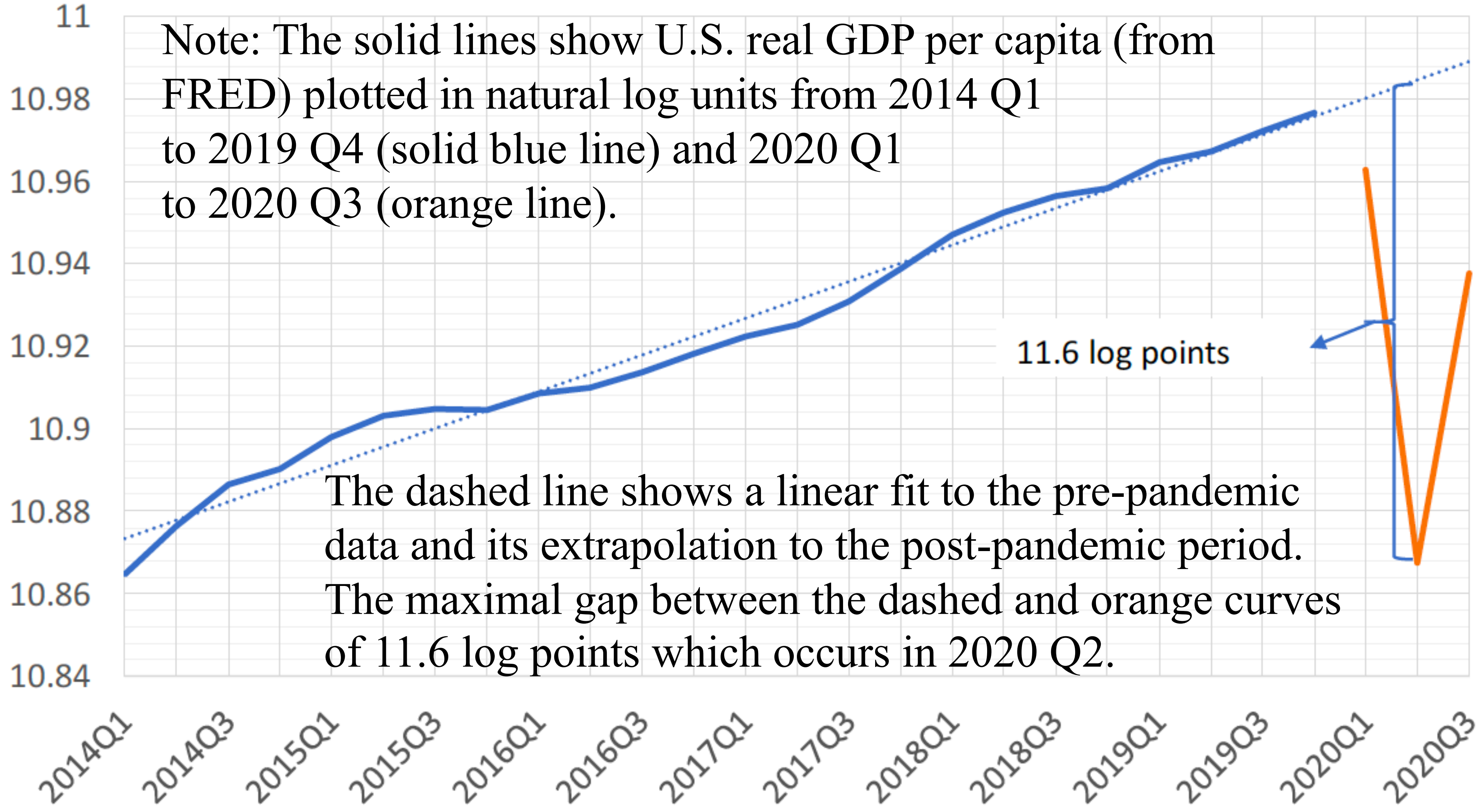
where ρ is the time preference rate, θ is relative risk aversion, σ is the standard deviation of the output growth rate absent disasters, E denotes the expectations operator, *p* is the disaster probability, and *b* is the size of the log output drop when disaster strikes. Agents know the parameters.

In taking this model to the data, we interpret 17 February as the last date before disaster strikes and 23 March as the date by which agents fully grasp the gravity of the disaster. Global and U.S. equity prices fell about 40 percent (51 log points) over this 33-day period. Using (3) and (4), the model-implied realized equity return over this period is

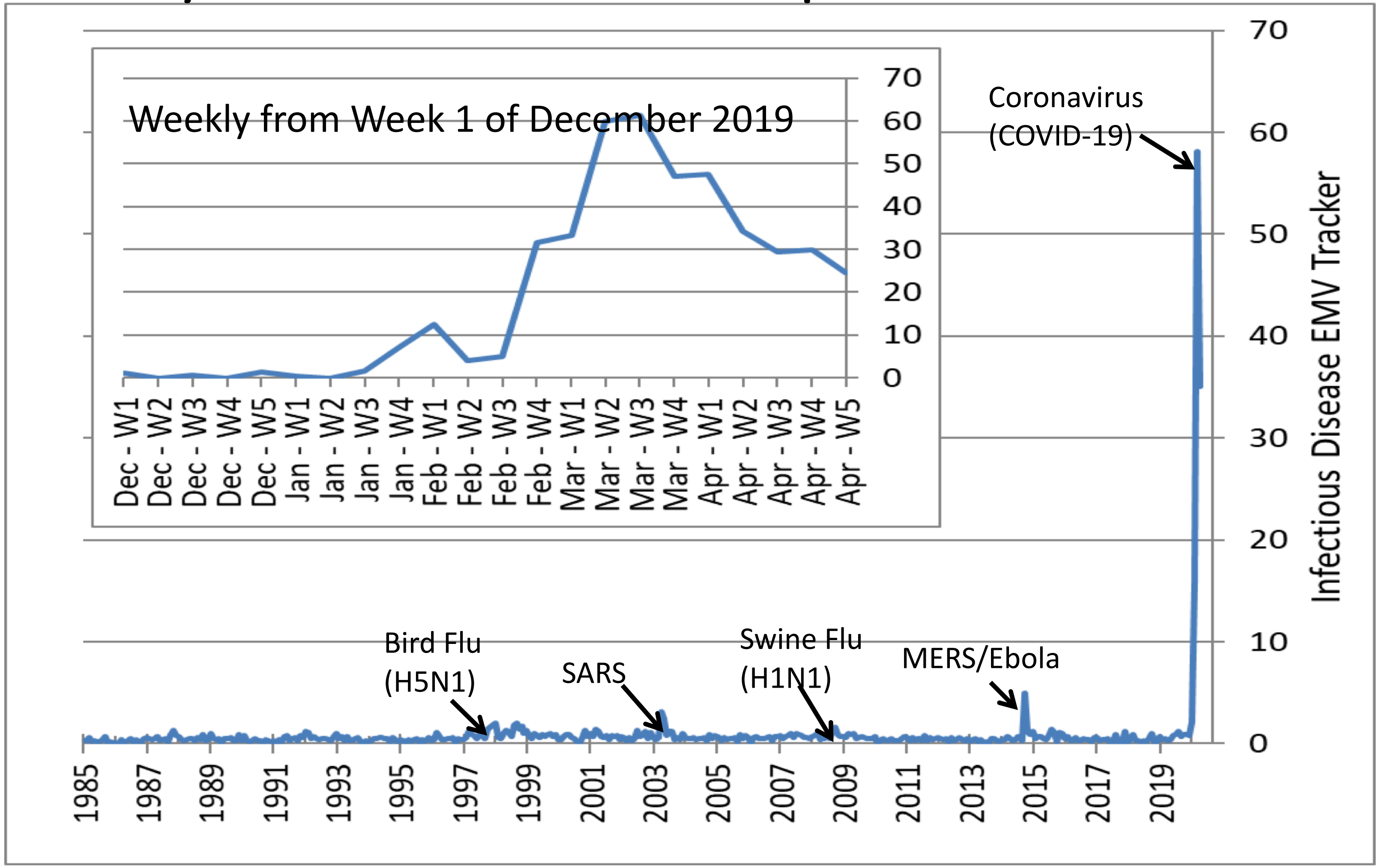
$$ln\left(\frac{P_{\text{after}}}{P_{\text{before}}}\right) = ln\left(\frac{A_{\text{after}}}{A_{\text{before}}}\right) = \gamma\left(\frac{33}{365}\right) + \mathrm{u}_1 - |\mathrm{v}_1|, \qquad (5)$$

where $|\mathrm{v}_1|$ is the realized disaster size, and u_1 is the realized value of the regular shock. For any reasonable values of the annual drift (γ) and the variability of regular shocks (σ), the first two terms on the right side are tiny compared to v_1. Thus, the model implies that stock prices should fall nearly one-for-one in proportion to disaster size. (Given the stochastic process in (3), the rates of return on one-period and full equity claims are identical.)

Assessing the Size of the COVID Disaster

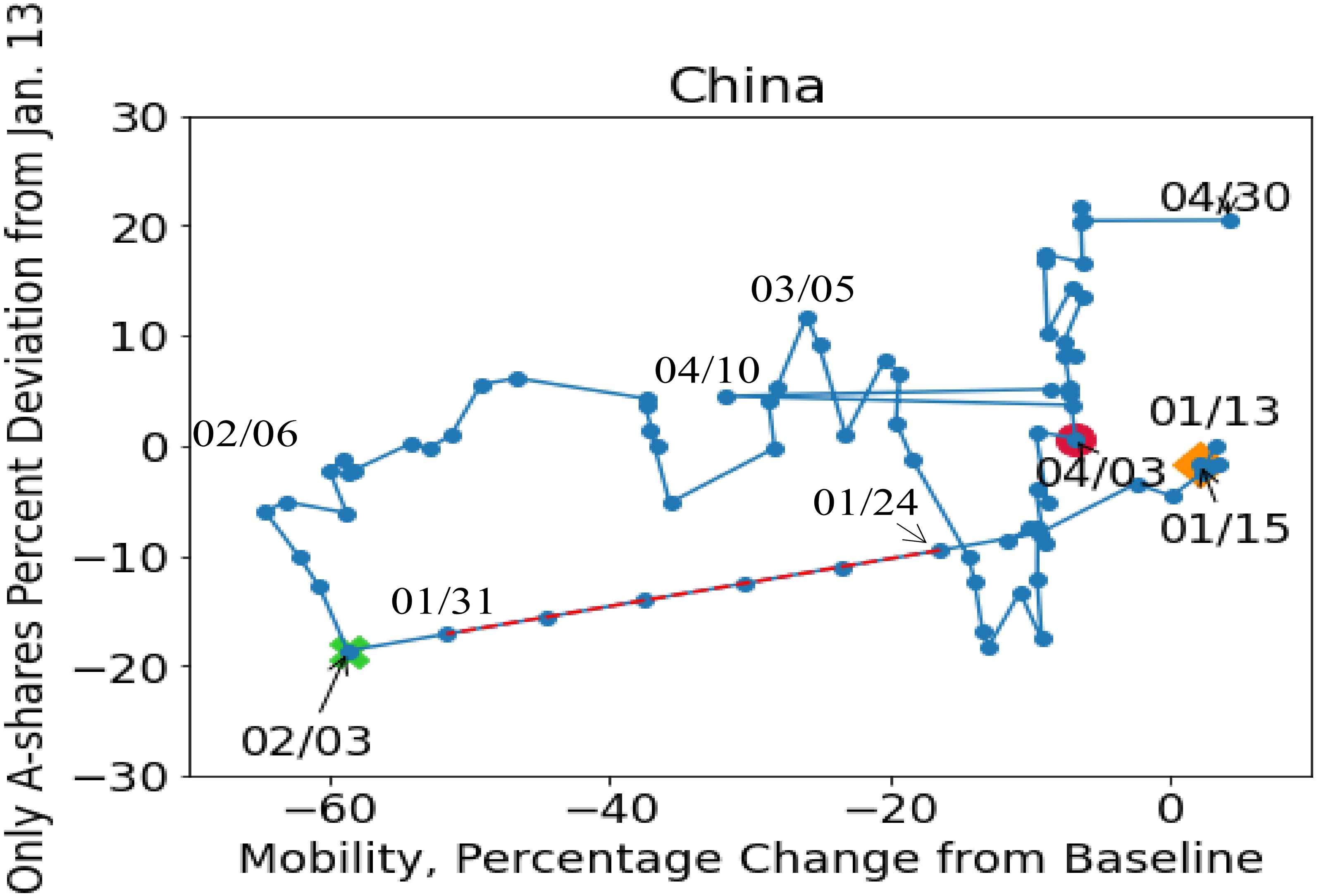


Infectious Disease EMV Tracker, Weekly and Monthly Data from 1985 to April 2020



Notes: The Infectious Disease EMV Tracker is computed as the overall Equity Market Volatility Tracker value multiplied by the share of EMV Articles that contain one or more of the following terms: epidemic, pandemic, virus, flu, disease, coronavirus, mers, sars, ebola., H5N1, H1N1. See Baker, Bloom, Davis and Kost (2019) and Baker, Bloom, Davis, Kost, Sammon and Viratyosin (2020).

Time Path of China Stock Prices and Mobility from 13 January to 30 April 2020



Note: Stock prices for companies with equity securities listed on mainland exchanges only and denominated in RMB from the CSMAR dataset (China analog to WRDS). An orange diamond marks the first confirmed COVID-19 death, a green cross marks the first date with stringency index value of 70 or more, and a red dot marks the date on which the stringency index first drops below 70. We linearly interpolate stock prices from 24 January to 3 February, given that mainland China stock markets were closed from 25 January to 2 February, inclusive.

Source: Davis, Liu and Sheng (2021).

The Unprecedented Stock Market Impact of the Coronavirus: China

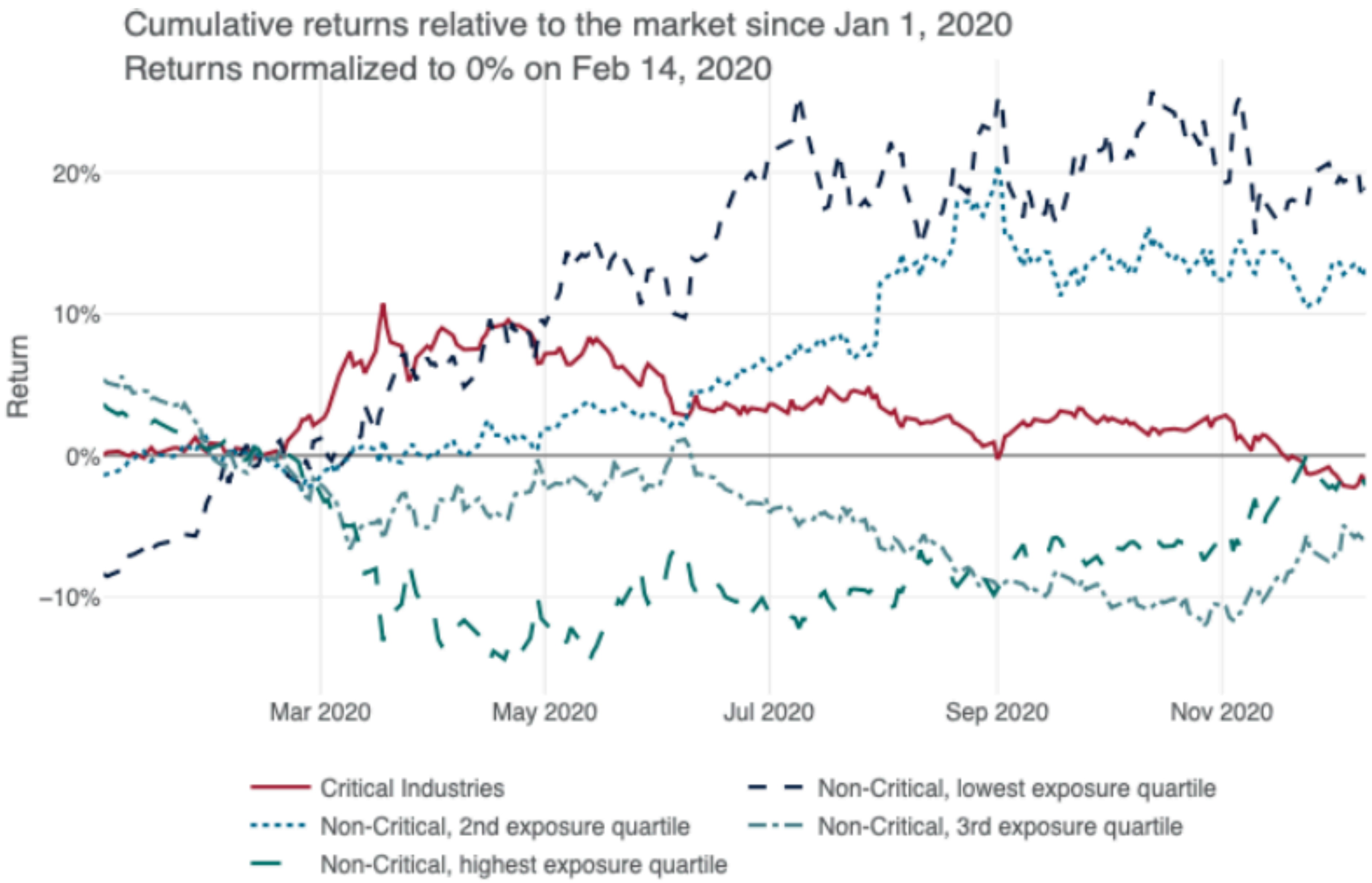
A. Shanghai Stock Exchange

Time Period	Jump Size	Number of Daily Stock Market Jumps	# Attributed to Economic Fallout of Pandemics	# Attributed to Policy Responses to Pandemics
26 December 1990 31 December 2019	≥ |4%|	384	0	0
2 January 2020 to 30 April 2020	≥ |4%|	1	1	0
	≥ |3%| and < |4%|	5	4	1

The same pattern holds for the Hang Seng index (Hong Kong).

Reproduced from “Stock Prices, Lockdowns, and Economic Activity in the Time of Coronavirus” by Davis, Liu and Sheng (2021).

Equity Markets Think the Shift to WFH Is a Big Deal



Firms outside ”Critical Industries” sorted into quartiles based on the fraction of workers in their industry that can feasibly work from home.

This chart is from https://sites.google.com/site/lawrencedwschmidt/covid19 and is based on work by Schmidt and Papanikalaou (2020).

Selected Firms with Big Fit Gains on Pandemic Fallout Days from MNIR

Company	Business description	Terms	tf-idf x MNIR coeff.
NOVAVAX INC ↗	Late-stage biotechnology company focused on the discovery, development and commercialization of vaccines to prevent serious infectious diseases.	vaccine influenza clinical trials candidates collaborators	475·2 297·5 136·9 99·4 96·1
NETFLIX INC ↗	World's leading internet television network with streaming memberships in over 190 countries.	dvd streaming subscribers titles studios	943·2 445·6 243·5 216·6 81·9
DOMINO'S PIZZA INC ↗	Multinational pizza restaurant chain with a large global network of franchise owners.	cheese quick foods pound interruption from earthquakes	56·5 28·4 25·5 20·9 16·8
PLAINS ALL AMER PIPELNE -LP ↘	Provider of midstream energy infrastructure and logistics services for crude oil, natural gas liquids, natural gas and refined products.	crude ngl barrels per day pipeline pipelines	−1639·7 −1499·2 −1363·3 −815·5 −681·7
MARCUS CORP ↘	Owner and operator of real estate assets in the lodging and entertainment industries: movie theatres, hotels and resorts, a family entertainment center.	films hotels movie film patrons	−216·8 −215·2 −164·6 −103·6 −72·2

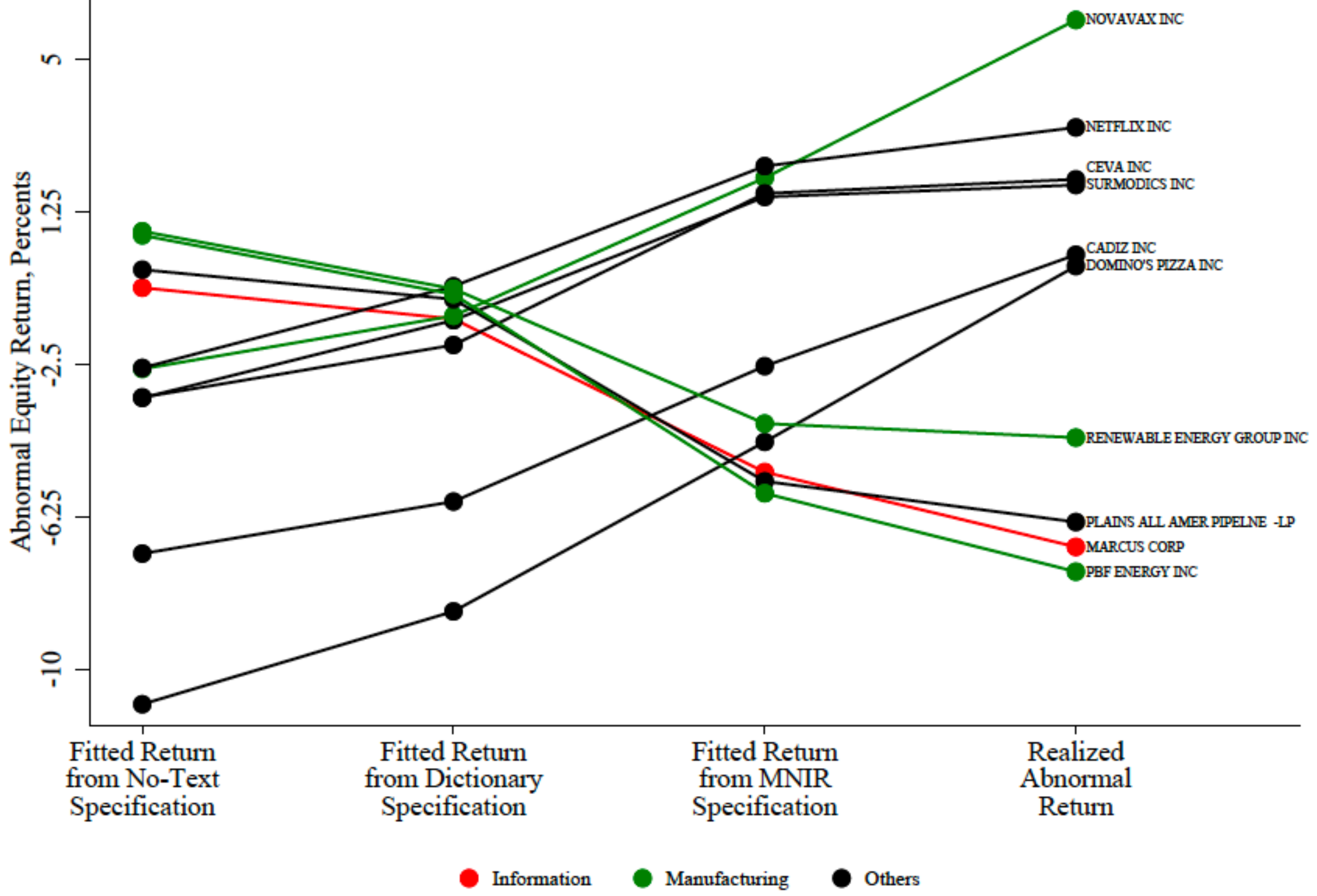
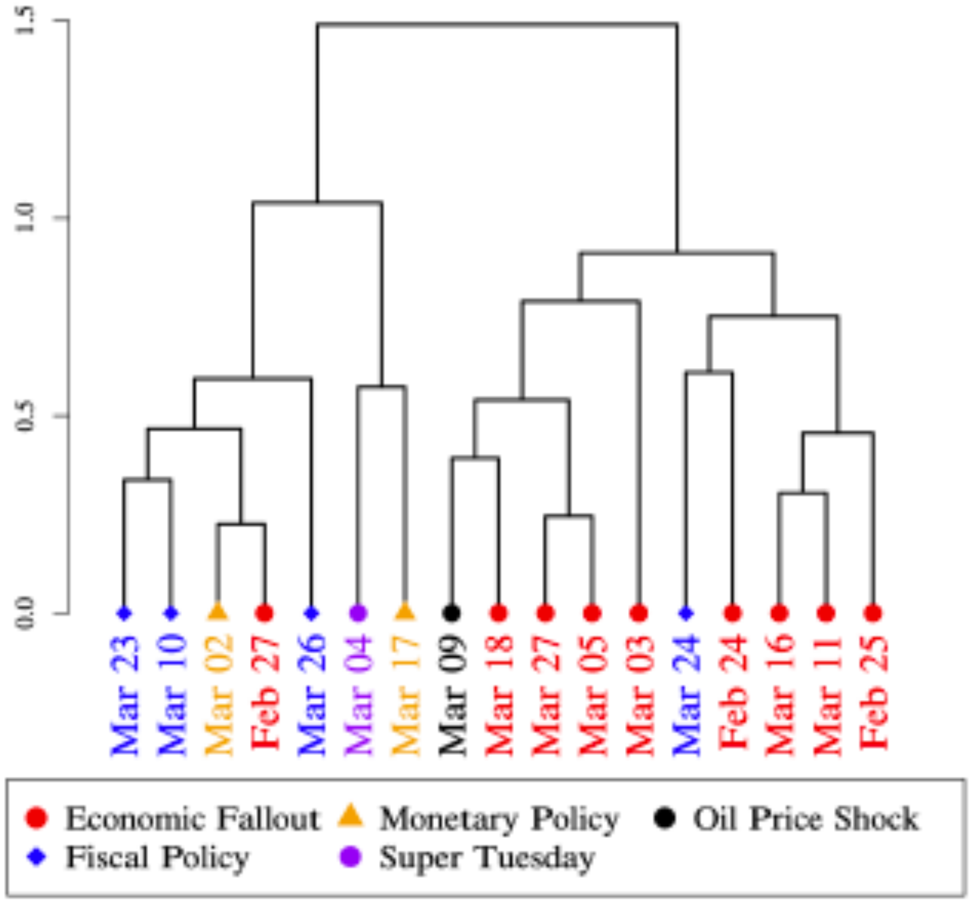


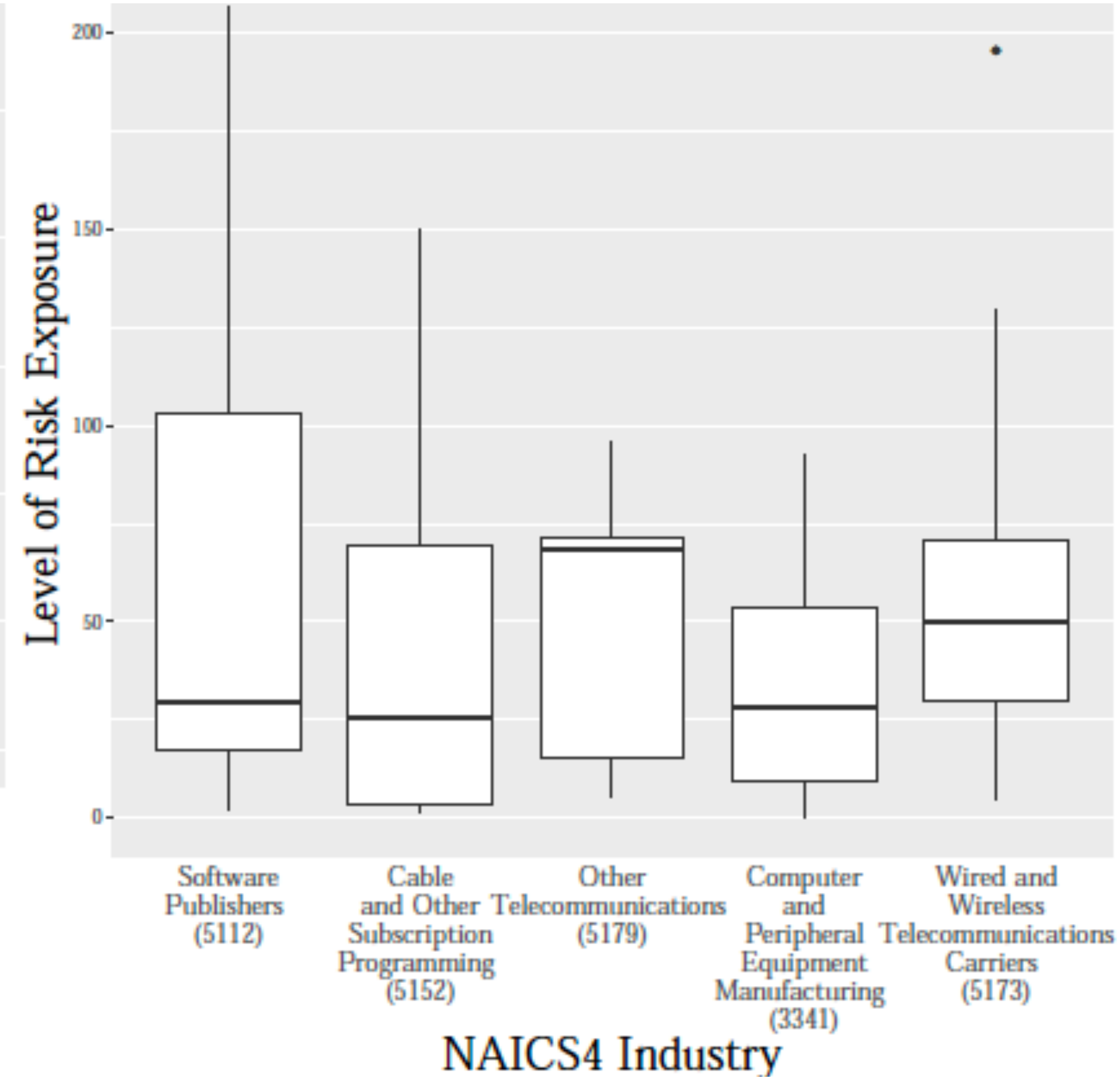
Figure B.8: Predicted and Actual Returns for Firms with Greatest MNIR Fit Gains on Pandemic Fallout Days

Clustering of Jump Days Based on MNIR-Fitted Structure of Returns to Classification Based on Human Readings of Next-Day Newspaper Explanations for Jumps



(b) Machine Learning Method

An Example of Intra-Industry Differences in Shock Exposures



This chart shows box plots for the distribution of firm-level exposures to “Web-Based Services” for the indicated industries.

References

Baker, Scott R., Nick Bloom, and Steven J. Davis, 2016. "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics.*

Baker, Scott R., Nick Bloom, Steven J. Davis and Kyle Kost, 2019. "Policy News and Stock Market Volatility," NBER WP 25720.

Baker, Scott R., Nick Bloom, Steven J. Davis, Kyle Kost, Marco Sammon and Tasaneeya Viratyosin, 2020. "The Unprecedented Stock Market Reaction to COVID-19," *Review of Asset Pricing Studies,* 10, no. 4.

Baker, Scott R., Nick Bloom, Steven J. Davis, and Marco Sammon and Tasaneeya Viratyosin, 2021. "What Triggers Stock Market Jumps?" NBER Working Paper 28687.

Barrero, Jose Maria, Nicholas Bloom and Steven J. Davis, 2020. "COVID-19 Is Also a Reallocation Shock," *Brookings Papers on Economic Activity,* Summer.

Barrero, Jose Maria, Nicholas Bloom and Steven J. Davis, 2021a. "Why Working from Home Will Stick," NBER Working Paper 28731.

Barrero, Jose Maria, Nicholas Bloom, Steven J. Davis and Brent Meyer, 2021b. "COVID-19 Is a Persistent Reallocation Shock," *AEA Papers & Proceedings,* May.

Barro, Robert J., 2006. "Rare disasters and asset markets in the twentieth century," *Quarterly Journal of Economics*, 121, no. 3.

Davis, Steven J. 2017. "Regularity Complexity and Policy Uncertainty: Headwinds of Our Own Making," working paper.

Davis, Steven J., Stephen Hansen and Cristhian Seminario-Amez, 2020. "Firm-Level Risk Exposures and Stock Price Reactions to COVID-19," NBER Working Paper 27867.

Davis, Steven J., Dingqian Liu and Xuguang Simon Sheng, 2021. "Stock Prices, Lockdowns, and Economic Activity in the Time of Coronavirus," NBER Working Paper 28320.

Lougran T. and B. Mcdonald, 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *Journal of Finance,* 66, no. 1.

Papanikolaou, Dimitris and Lawrence D.W. Schmidt, 2020. "Working Remotely and the Supply-Side Impact of COVID-19," NBER Working Paper 27330.

Taddy, Matt, 2013. "Multinomial Inverse Regression for Text Analysis," *Journal of the American Statistical Association,* 1108, no 503.

Taddy, Matt, 2015. "Distributed Multinomial Regression," *Annals of Statistics,* 9, no. 3, 1394-1414.

Tetlock, Paul C., 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *Journal of Finance,* 62. no. 3.