Macro Shocks and Firm-Level Response Heterogeneity

Steven J Davis,[1] Stephen Hansen,[2] and Cristhian Seminario-Amez[3]

[1]Hoover Institution

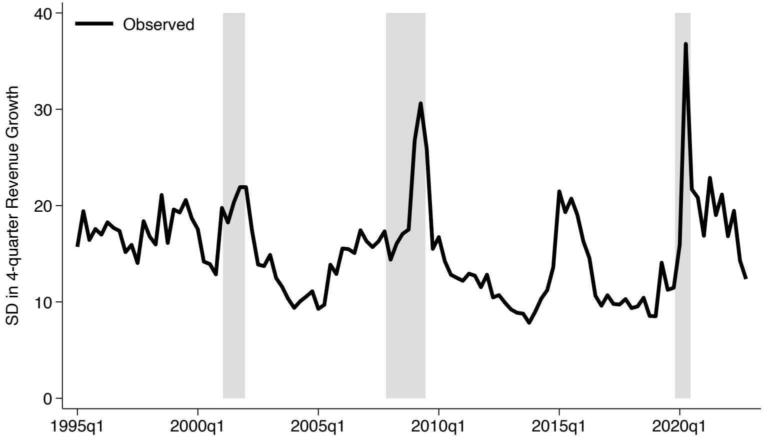
[2]UCL, IFS, and CEPR

[3]ITAM

December 2025

Outline

Firm-Level Outcomes are Countercyclical



Source: Balanced panel of firms from COMPUSTAT

Existing Explanations

Large literature (Bloom 2009, Bloom et. al. 2018, Senga 2018) posits that the variance of idiosyncratic firm-level shocks is countercyclical.

In Bloom et. al. 2018 firm j output in time t is

$$y_{j,t} = A_t z_{j,t} f(k_{j,t}, n_{j,t})$$

where $\log(A_t)$ and $\log(z_{j,t})$ are AR(1) processes with stochastic volatility.

Existing Explanations

Large literature (Bloom 2009, Bloom et. al. 2018, Senga 2018) posits that the variance of idiosyncratic firm-level shocks is countercyclical.

In Bloom et. al. 2018 firm j output in time t is

$$y_{j,t} = A_t z_{j,t} f(k_{j,t}, n_{j,t})$$

where $\log(A_t)$ and $\log(z_{j,t})$ are AR(1) processes with stochastic volatility.

Another view (Bachmann and Moscarini 2012, Ilut et. al. 2018, Berger and Vavra 2019) is that firms react more strongly to negative than to positive macro shocks.

Our Explanation

1. Firms are heterogeneously exposed to macro shocks. For example:

$$y_{j,t} = A_t^{\theta_j} z_{j,t} f(k_{j,t}, n_{j,t})$$

where θ_j is a firm-level exposure to macro conditions.

Our Explanation

1. Firms are heterogeneously exposed to macro shocks. For example:

$$y_{j,t} = A_t^{\theta_j} z_{j,t} f(k_{j,t}, n_{j,t})$$

   where θ_j is a firm-level exposure to macro conditions.

2. Macro shocks, and exposures to them, are distinct. For example:

$$y_{j,t} = A_{t,1}^{\theta_{j,1}} A_{t,2}^{\theta_{j,2}} z_{j,t} f(k_{j,t}, n_{j,t})$$

   where $\theta_{j,s}$ is a firm-level exposure to macro shock s.

Measuring Shock Exposure

How to isolate effect of specific shock exposure on real outcomes?

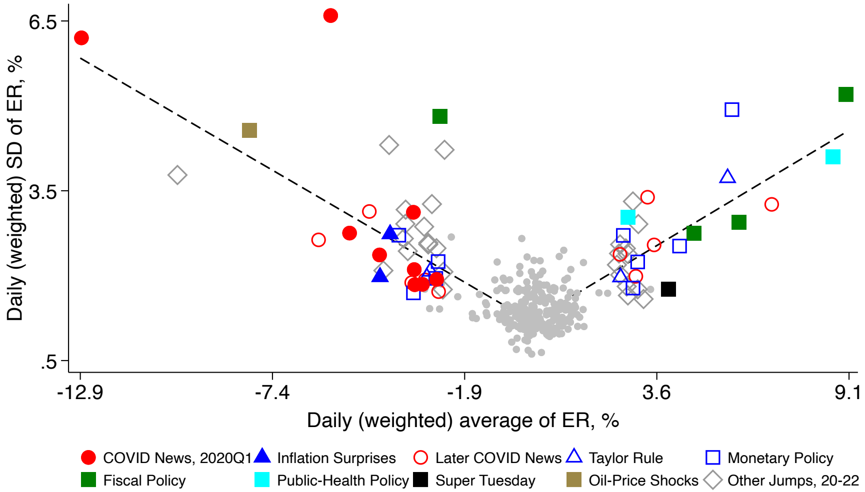
Step 1

Use high-frequency changes in firm-level stock prices in response to news about specific shocks (Kuttner 2001).

Days with macro news shocks are "jump dates" with aggregate market move $> 2.5\%$ (Baker et. al. 2025).

Jump dates are grouped together based on source of triggering news according to next-day newspaper accounts.

Return Dispersion on Jump Dates (2020-2022)



Measuring Shock Exposure

How to isolate effect of specific shock exposure on real outcomes?

Step 1

Use high-frequency changes in firm-level stock prices in response to news about specific shocks (Kuttner 2001).

Days with macro news shocks are "jump dates" with aggregate market move $> 2.5\%$ (Baker et. al. 2025).

Jump dates are grouped together based on source of triggering news according to next-day newspaper accounts.

Step 2

Problems with raw returns: (i) potentially vary for non-fundamental reasons; (ii) no interpretation of heterogeneous exposure.

We seek to isolate part of return explained by pre-shock firm characteristics.

Example

Company	NAICS	COVID News Return[1]	Rev Growth 18Q3-20Q3
NetApp Inc	3341[2]	4.34%	-6.9%
Scientific Games	3341	-4.92%	-64.2%
Domino's Pizza	722511[3]	-0.04%	20.8%
Ruth's Hospitality Group	722511	-9.28%	-44.5%

Different outcomes arise from characteristics not captured by sector membership (or other standard controls).

[1] Average return on "COVID News, 2020Q1" dates as defined above
[2] Computer and Peripheral Equipment Manufacturing
[3] Full-Service Restaurants

How to Measure Pre-Existing Business Characteristics?

To measure pre-existing characteristics, we use the Risk Factors discussions of 10-K filings.

RF texts detail sources of risk/uncertainty in future earnings; exhaustive due to their legal status.

Full set of RF corpora downloaded for the 2015-2019 time period. We use one compiled report by firm.

Matched to stock return data to form sample of ca. 2,000 firms.

Built text-based shock exposure using sufficient reduction projection from multinomial inverse regression (Taddy 2013, Taddy 2015, Gentzkow et. al. 2019, Cage et. al. 2021).

Main Findings

Our text-based exposures:

1. Predict real outcomes (revenue, investment, employment growth).

   Unexplained variation in returns is uncorrelated with real outcomes.

2. Provide an interpretable account of heterogeneity.

   Highlight supply-chain linkages, technology adoption, product characteristics, among others.

3. Account for most of the rise in revenue growth dispersion during the 2020 recession.

Outline

Quantifying Firm-Level Shock Transmission

Main real outcome of interest is quarterly revenue growth

$$\Delta \mathrm{rev}_{it} = \log(\mathrm{rev}_{i,t}) - \log(\mathrm{rev}_{i,t-12}) \quad t = 2018Q1, \ldots, 2022Q4$$

Balanced panel built from COMPUSTAT $\approx$ 1,000 firms.

Heterogeneous shock exposure $\Longrightarrow$ heterogeneous real outcomes?

Baseline Regression Model

$$\Delta \mathrm{rev}_{it} = I_i + I_{s(i),t} + \sum_{t=19Q1}^{22Q4} I_t \boldsymbol{\alpha}_t' \mathbf{e}_i + \sum_{t=18Q1}^{18Q4} I_t \boldsymbol{\beta}' \mathbf{controls}_{it} + \sum_{t=19Q1}^{22Q4} I_t \boldsymbol{\beta}_t' \mathbf{controls}_{it} + \epsilon_{it}$$

- Firm fixed effects

Baseline Regression Model

$$\Delta\mathrm{rev}_{it} = I_i + I_{s(i),t} + \sum_{t=19Q1}^{22Q4} I_t \alpha_t' \mathbf{e}_i + \sum_{t=18Q1}^{18Q4} I_t \beta' \mathbf{controls}_{it} + \sum_{t=19Q1}^{22Q4} I_t \beta_t' \mathbf{controls}_{it} + \epsilon_{it}$$

- Firm fixed effects
- Shock exposure measures

Baseline Regression Model

$$\Delta \mathrm{rev}_{it} = I_i + I_{s(i),t} + \sum_{t=19Q1}^{22Q4} I_t \boldsymbol{\alpha}_t{}' \mathbf{e}_i + \sum_{t=18Q1}^{18Q4} I_t \beta' \textbf{controls}_{it} + \sum_{t=19Q1}^{22Q4} I_t \beta_t' \textbf{controls}_{it} + \epsilon_{it}$$

- Firm fixed effects.
- Shock exposure measures.
- Key coefficients of interest $\longrightarrow$

  Does variation in exposure shift revenue growth relative to 2018?

$$\mathbf{e}_i = \left(\mathrm{AbnRet}_i^P, \mathrm{AbnRet}_i^I\right)$$

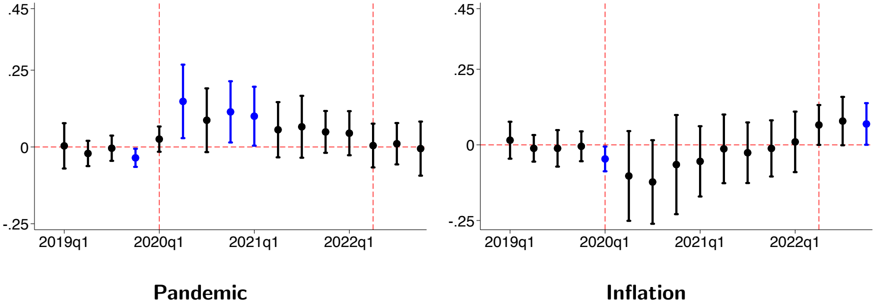


Figure: Revenue Growth and **Raw-Return** Exposures

Outline

Bag-of-Words Representation

All 10-K files are pre-processed using standard steps, converts each firm's *RF* filings into list of terms.

$\approx$ 16,000 **unique** terms and $\approx$ 45,000,000 **total** terms.

Let $c_{i,v}$ be the count of term v in firm i's *RF* filings and $C_i = \sum_v c_{i,v}$ the firm-level total count.

Forms document-term matrix with high-dimensional column space. $\mathbf{c}_i$ is the ith row.

Multinomial Inverse Regression (Taddy 2015)

Data generating process is multinomial logit

$$\mathbf{c}_i \sim \mathrm{MN}(\mathbf{q}_i^s, C_i) \quad \text{where} \quad q_{i,v}^s = \frac{\exp(\eta_{i,v}^s)}{\sum_v \exp(\eta_{i,v}^s)}$$

parametrized as

$$\eta_{i,v}^s = a_v^s + (\mathbf{x}_i^s)^T \mathbf{b}_v^s$$

with covariates $\mathbf{x}_i^s = (AbnRet_i^s, \mathrm{NAICS2}_i, \mathrm{size}_i, \mathrm{leverage}_i)$

Palmgren (1981) shows that MLE estimates for coefficients correspond to those under an alternative model in which $c_{i,v}$ are drawn independently as

$$c_{i,v} \sim \mathrm{Poisson}(\exp(\mu_i^s + \eta_{i,v}^s)).$$

Likelihood Approximation

Poisson log-likelihood kernel is

$$\sum_i \sum_v [c_{i,v}(\mu_i + \eta_{i,v}) - \exp(\mu_i + \eta_{i,v})]$$

Taddy (2015) suggests the plugin estimate $\hat{\mu}_i = \log(C_i)$ under which the above simplifies to

$$\sum_i \sum_v [c_{i,v}\eta_{i,v} - C_i \exp(\eta_{i,v})]$$

This linearly separates the regression coefficients for each term.

Permits the fitting of V Poisson regressions in parallel; highly efficient.

Each Poisson regression is fit with a separate LASSO penalty where strength determined by AIC.

Ranking of Terms for Pandemic Dates

Bottom Terms		Top Terms	
Term	**Value**	**Term**	**Value**
hotel	-33549.80	game	19595.17
unitholder	-18019.33	product_candidate	18475.38
general_partner	-14038.83	client	17157.35
gaming	-11840.90	drug_candidate	12859.84
restaurant	-11830.94	clinical_trial	11963.56
reit	-10605.63	cellular	9029.67
tenant	-10177.66	subscription	7706.73
satellite	-9343.81	solution	7571.91
crude_oil	-9220.05	patient	6222.35
common_unit	-9087.59	student	6058.91
hotel_property	-8337.83	platform	5430.79
refinery	-7749.49	drug	5120.28
travel	-6931.05	collaborator	4641.59
franchisee	-6722.54	datum_center	4317.81
trs	-6657.09	celgene	4266.25

Table: **Influential Terms for Pandemic Jumps**. The ranking of term v is based on its tf-idf score multiplied by $\hat{\beta}_v^P$.

Ranking of Terms for Inflation Dates

Bottom Terms		Top Terms	
Term	**Value**	**Term**	**Value**
tenant	-15810.95	hotel	31389.56
student	-9557.59	natural_gas	9899.49
operating_partnership	-7896.48	gaming	8074.73
reit	-7610.02	crude_oil	8016.95
real_estate	-7081.87	hotel_property	7193.75
homebuilding	-7078.13	aircraft	7111.44
the_company	-7038.95	solar	7065.24
product_candidate	-6580.21	pipeline	6868.64
home	-6515.21	oil	6852.39
client	-6057.50	ferc	6846.26
property	-5597.77	unitholder	6420.11
cellular	-5487.07	drilling	6027.62
fcc	-5200.31	ngl	5968.09
clinical_trial	-4671.19	fuel	5235.03
wireless	-4270.11	semiconductor	5142.39

Table: **Influential Terms for Inflation Jumps**. The ranking of term v is based on its tf-idf score multiplied by $\hat{\beta}^I_v$.

Distinct Relationship among Terms and Returns

Sign of Terms in Inflation Regression Sign of Terms in Pandemic Regression	Negative	Zero	Positive
Negative	1701	1289	1833
Zero	1288	1220	1027
Positive	2243	1516	1713

Table: **Cross Tabulation**. Tabulation of the signs of each 10-K Risk Factors term in the inverse regressions associated with pandemic fallout and inflation jump days. LASSO penalization generates a zero coefficient for some terms.

Sufficient Reduction Projection (Cook 2007)

We can write the Poisson pmf for $c_{i,v}$ in exponential form as

$$p(c_{i,v} \mid \mathbf{x}_i) = \frac{1}{c_{i,v}!} \exp\left[c_{i,v}(\hat{\mu}_i + \eta_{i,v}) - \exp(\hat{\mu}_i + \eta_{i,v})\right]$$

so that that the probability of the full count vector is

$$p(\mathbf{c}_i \mid \mathbf{x}_i) = h(\mathbf{c}_i) \exp\left[\sum_v c_{i,v} a_v + \sum_j \underbrace{\left(\sum_v c_{i,v} b_{j,v}\right)}_{z_{i,j}} x_{i,j} + g(\boldsymbol{\eta}_i)\right]$$

$z_{i,j}$ is a sufficient reduction projection for the jth variable in $\mathbf{x}_i$; focus on SRP for AbnRet_i^s which we denote z_i^s.

$p(\mathrm{AbnRet}_i^s \mid \mathbf{c}_i, \mathbf{x}_i)$ only depends on the high-dimensional count vector $\mathbf{c}_i$ via the scalar z_i^s.

Shock Exposures are Distinct

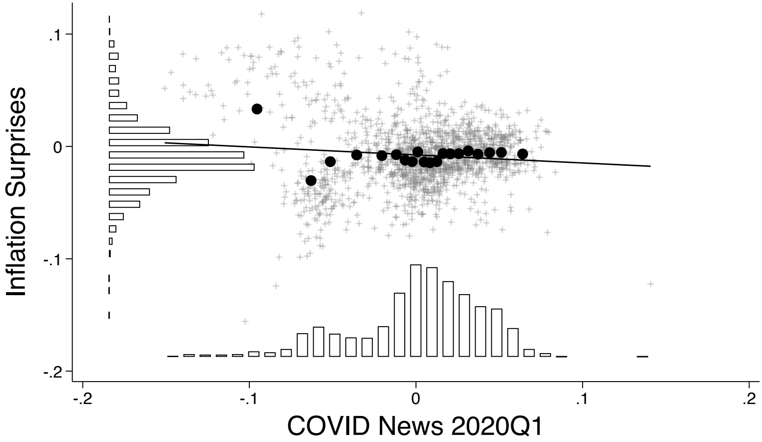


Figure: Distributions of Pandemic and Inflation Exposures

Exposure Components

Each exposure is composed of thousands of individual terms which obscures source of heterogeneity.

We propose a clustering algorithm that combines terms for each jump date that depends on (i) common sign in inverse regression (ii) semantic similarity from word embedding model.

Examples:

Negative Pandemic Component 11: *airline, airport, travel, flight, passenger, ticket, route, air travel*

Negative Inflation Component 2: *home, housing, new home, homebuyer, mortgage financing, mortgage interest, availability of mortgage*

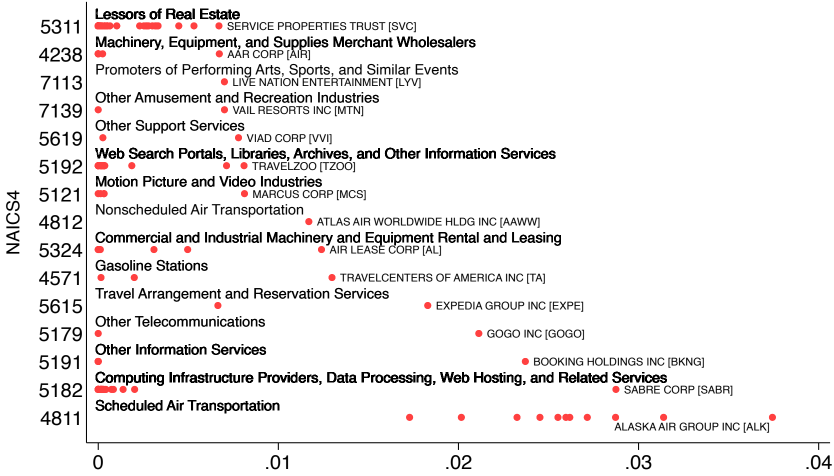


Figure: **Distribution of 'Travel' Pandemic Exposure**
We choose the top ten NAICS4 industries ranked according to maximally exposed firm. Each red dot represents a firm.

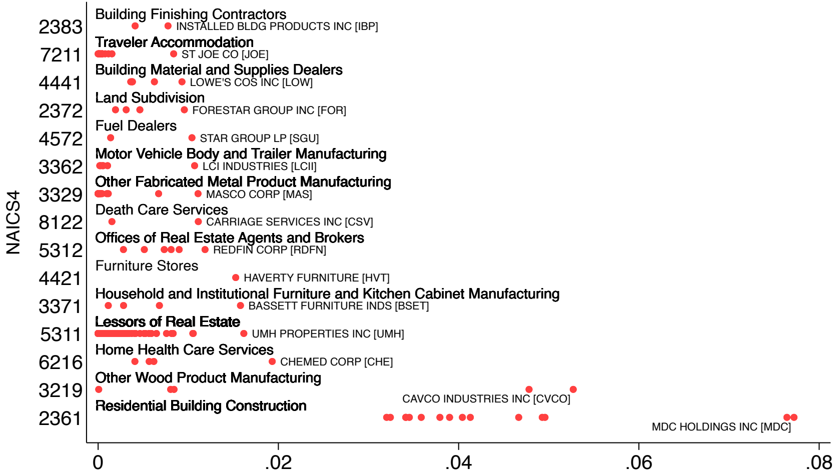


Figure: **Distribution of 'Mortgages and Real Estate' Inflation Exposure**
We choose the top ten NAICS4 industries ranked according to maximally exposed firm. Each red dot represents a firm.

Outline

Returns Regressions

$$\mathrm{AbnRet}_i^s = \alpha_0^s + \left(\boldsymbol{\alpha}_1^s\right)^T \mathbf{x}_i^{y(s)} + \epsilon_i^s$$

$$\mathrm{AbnRet}_i^s = \gamma_0^s + \left(\boldsymbol{\gamma}_1^s\right)^T \mathbf{x}_i^{y(s)} + \gamma_2^s z_i^s + \gamma_3^s C_i + \varepsilon_i^s$$

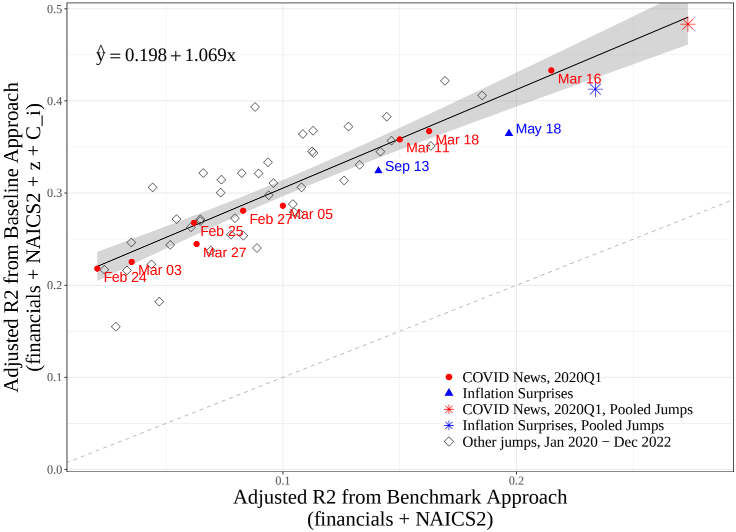


Figure: Adjusted R^2 from regressing abnormal returns from jump dates on controls for NAICS2 sector, leverage, and size (x-axis) and additionally including text-based shock exposures and controls for text length (y-axis).

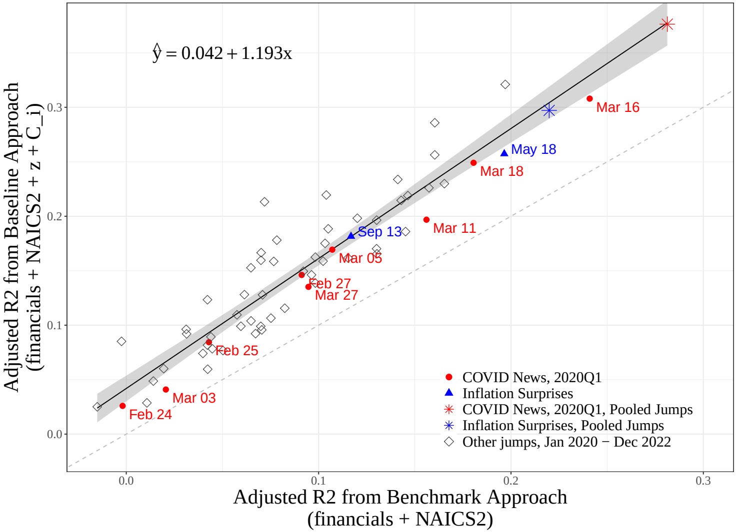


Figure: Adjusted R^2 from regressing abnormal returns from jump dates on controls for NAICS2 sector, leverage, and size (x-axis) and additionally including text-based shock exposures and controls for text length (y-axis).

Out-of-Sample Dates

	(1) Second Wave Fears (2020/06/11)	(2) Vaccine News (2020/05/18)	(3) Slower Fed Tightening (2022/11/10)
Exposure Variable	z_i^P	z_i^P	z_i^I
Coefficient	0.51***	-0.46***	-0.23***
	(0.032)	(0.060)	(0.026)
Observations	1,552	1,552	1,506
Adjusted R^2	0.349	0.339	0.116
Adjusted R^2 (only controls)	0.199	0.207	0.069

Outline

Review: $\mathbf{e}_i = (\mathrm{AbnRet}_i^P, \mathrm{AbnRet}_i^I)$

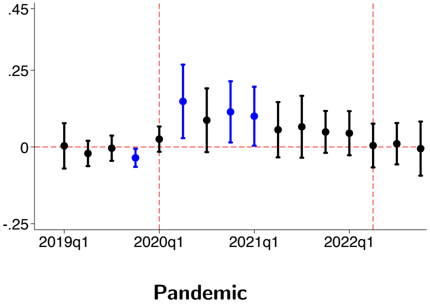
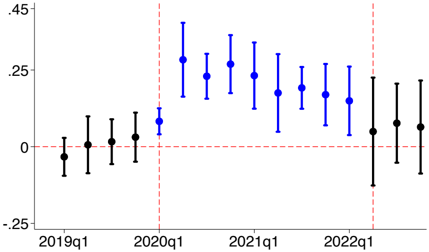
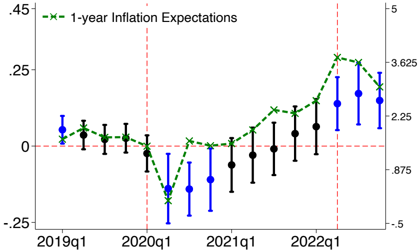


Figure: Revenue Growth and **Raw-Return** Exposures

New Specification: $\mathbf{e}_i = (z_i^P, \hat{\varepsilon}_i^P, z_i^I, \hat{\varepsilon}_i^I)$



Pandemic Text Exposure



Inflation Text Exposure

Figure: Revenue Growth and **Text** Exposures

New Specification: $\mathbf{e}_i = \left(z_i^P, \hat{\varepsilon}_i^P, z_i^I, \hat{\varepsilon}_i^I\right)$

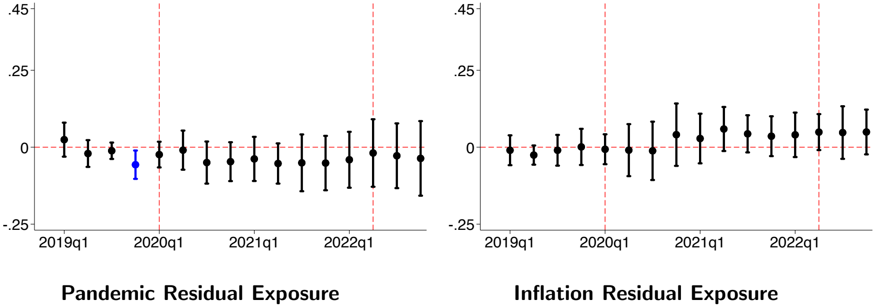


Figure: Revenue Growth and **Residual** Exposures

Additional Results

Similar findings for (annual) growth in revenue, employment, and investment.

Text exposures explain more variation in 2020Q3 earnings surprises than abnormal returns.

2020Q1 exposure measures from Hassan et. al. (2023) have limited forward-looking information.

Outline

Shock Exposures Account for Much of Realized Dispersion

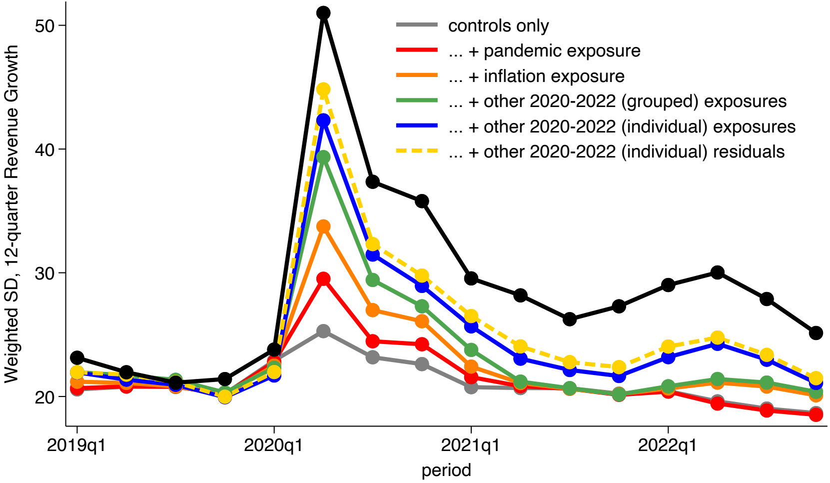
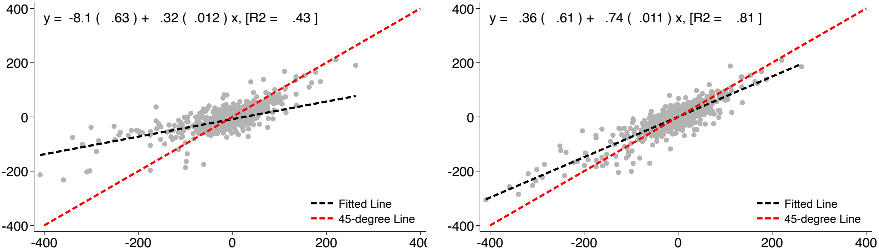


Figure: Realized and Fitted Standard Deviation of Revenue Growth

Distribution of 2020Q2 Revenue Growth



Firm FE + Controls **+ All Text Exposures**

Figure: Realized and Predicted Revenue Growth in 2020Q2

Historical Evidence I

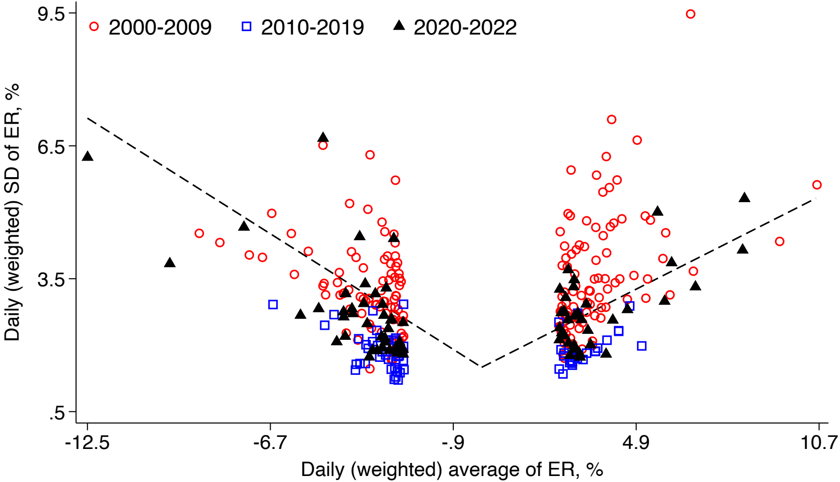


Figure: Abnormal Return Dispersion on Jump Dates (2000-2022)

Historical Evidence II

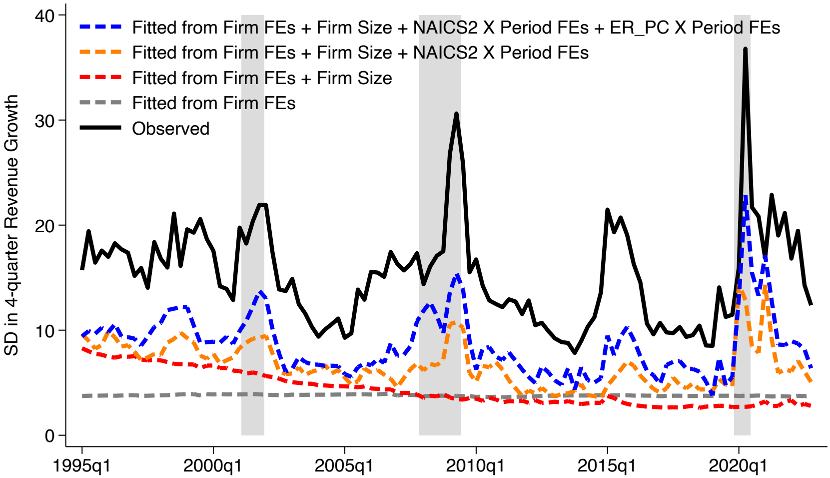


Figure: Dispersion in Firm-Level Revenue Growth since 1995

Outline

1. Motivation
2. Shock Exposures and Real Outcomes: Returns-Based
3. Constructing Firm-Level Exposures with Text
4. Text and Return Responses
5. Shock Exposures and Real Outcomes: Text-Based
6. Shock Exposures and Distribution of Firm-Level Outcomes
7. Conclusion

Conclusion

We introduce a method for building firm-level exposures to macro shocks via rich textual descriptions of pre-shock business characteristics.

These exposures explain much of the increase in firm-level growth dispersion in the 2020 recession.

Research implications for quantitative models of firm dynamics + de-noising asset-based measures of shock exposures.

Policy implications for targeted and shock-specific support.

Work in Progress: Tariff Shocks