Price Revelation from Insider Trades: Evidence from Hacked Earnings News

Lecture 13 MATH 60230 (review session)

Pat Akey
University of Toronto

Vincent Grégoire
HEC Montréal

Charles Martineau
University of Toronto

February 2026

One of the largest securities fraud cases in US history

From 2010 to 2015, a group of Ukranian hackers breached the IT systems of 3 of the largest newswire companies.

- Accessed earnings press releases several hours before their scheduled release.
- Sold the information to a select group of traders.

Traders aggressively traded before the news was publicly released to exploit this private information.

In 2015, the SEC charged some of the traders based in the U.S. for illegal insider trading for profits amounting to +$100 million.

Research Questions

This setting allows us to answer the following questions:

1. How do the fundamental characteristics of stocks or the characteristics of private signals affect an informed investor's choice of trading strategy?
2. (How) do prices incorporate private information imbedded in trades?
3. Can market makers respond to increased informed trading?

Our setting allows us to compare the price discovery dynamics of a group of "treated" firms whose earnings were exposed at a particular point in time to a "control" set whose earnings were not exposed at that point in time

- Hackers intermittently gained and lost access to newswires' IT systems at different points in time.
  - Will allow us to conduct a difference-in-differences analysis.

Preview of results

For stocks exposed to hacks:

1. Hackers more likely to trade when firms were larger, had larger analyst coverage, and had lower spreads, as well as for earnings news that had both hard and soft signals pointing in the same direction
2. Overnight returns are 15% less responsive to earnings surprises or soft information
   - Evidence that price discovery occurs before the closing of markets

Increased informed trading caused liquidity providers to increase spreads

1. Equity volume, option volume, equity turnover were higher
2. Increase *informed* order flow associated with higher effective and realised spreads (not in paper)
   - Not true for unconditional order flow

Empirical Design

Problems that confront empirical studies of insider trading detection:

- Identifying counterfactual trades that *could have had informed trading* is difficult
  - Recent papers focusing on prosecuted cases (e.g., Ahern (2020), Kacperczyk and Pagnotta (2020)) often rely on insider trading before mergers (52% of cases) — what is the counterfactual?
  - Collin-Dufresne and Fos (2018) use time periods where institutions take large positions in firms, assuming this is informed but do not identify the type of information
- Unobserved heterogeneity either at the firm level, or in the trading environment will likely lead to biased inference
- In equilibrium, informed traders likely strategically time their behavior so as to minimize detection

Our setting allows us to observe plausibly exogenous variation in the information set of informed traders around a common set of events with a limited ability to engage in strategic timing

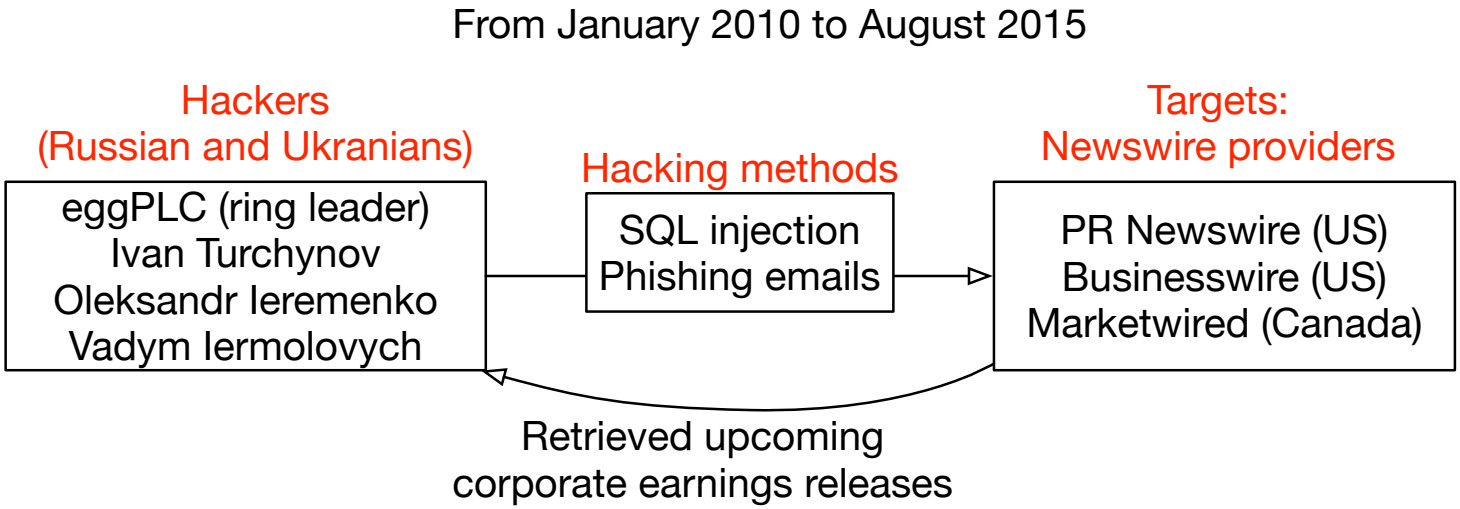
The hacking and insider trading scheme

From January 2010 to August 2015

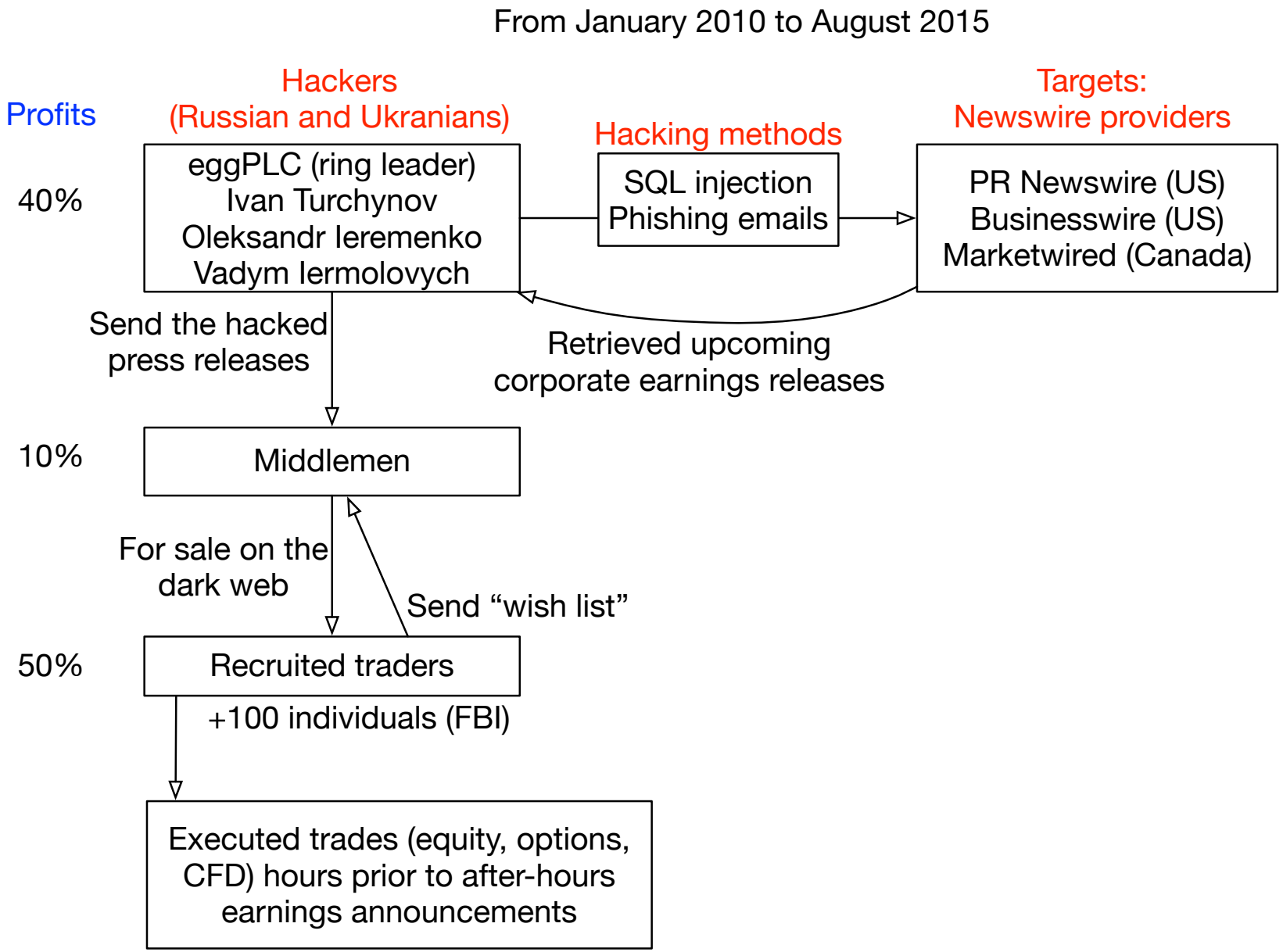
Hackers
(Russian and Ukranians)

eggPLC (ring leader)
Ivan Turchynov
Oleksandr Ieremenko
Vadym Iermolovych

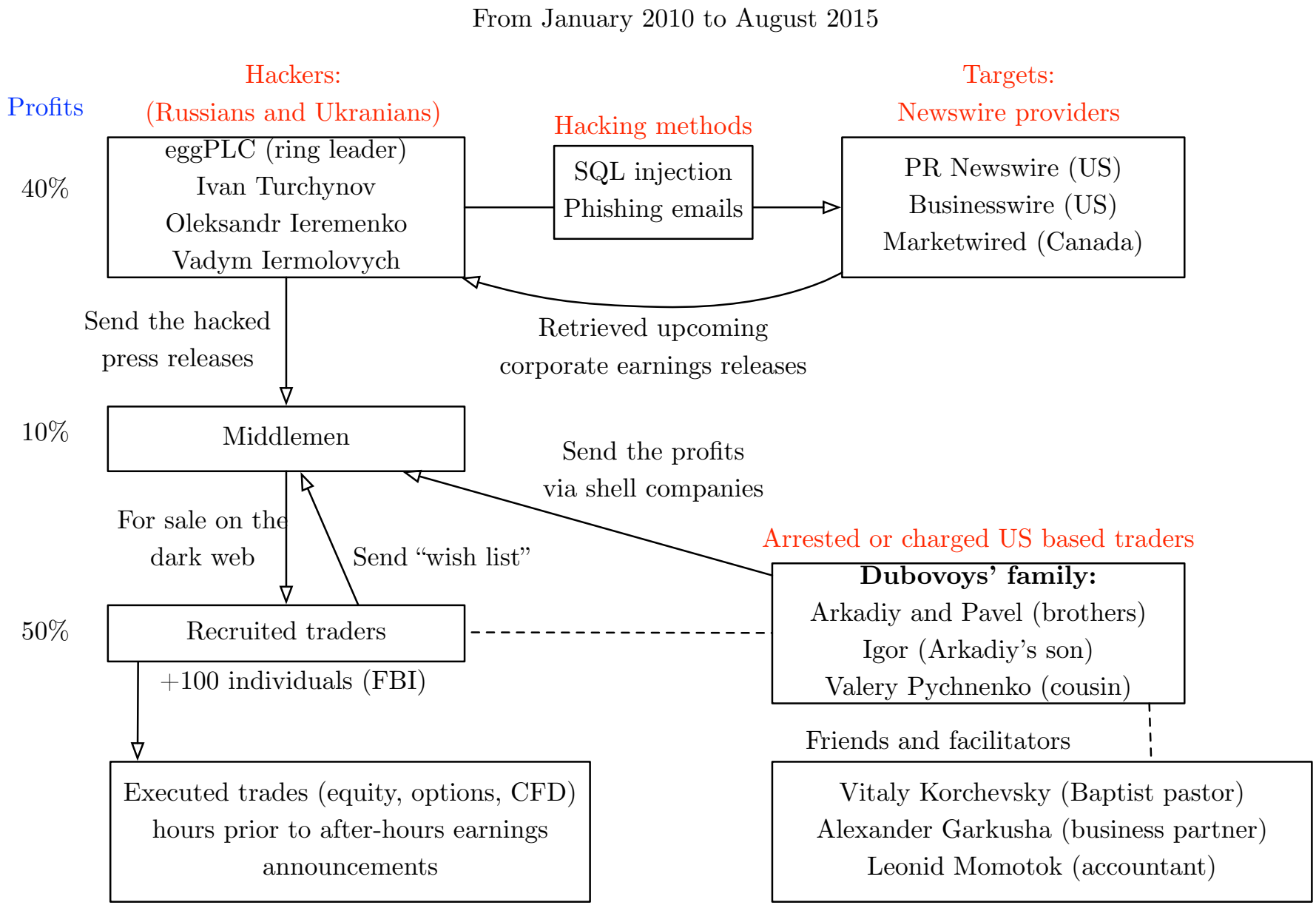
The hacking and insider trading scheme



The hacking and insider trading scheme



The hacking and insider trading scheme



How they were caught ...

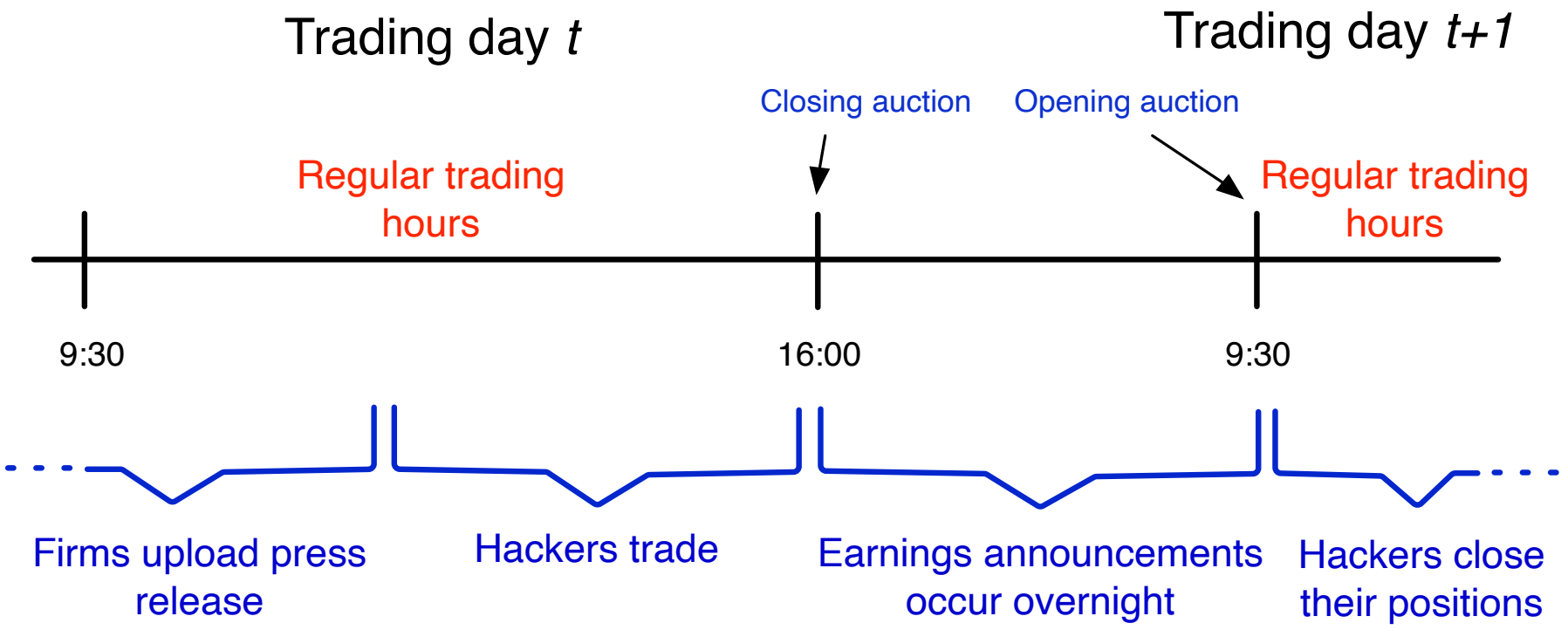
$100 million represents only a fraction of the money authorities believe was made off the stolen press releases.

As of August 2015, 20 individual traders have been charged.

Factors leading to the prosecutions:

- The arrest of the hacker Vadym Iermolovych in Mexico.
- The SEC, with the help of FINRA, developed algorithms to detect stock price fluctuations caused by some trades before corporate announcements and investigate the entity related to the flagged trades.
- Not an easy task since insiders used multiple accounts but the owners of the accounts were linked similar social or familial networks.

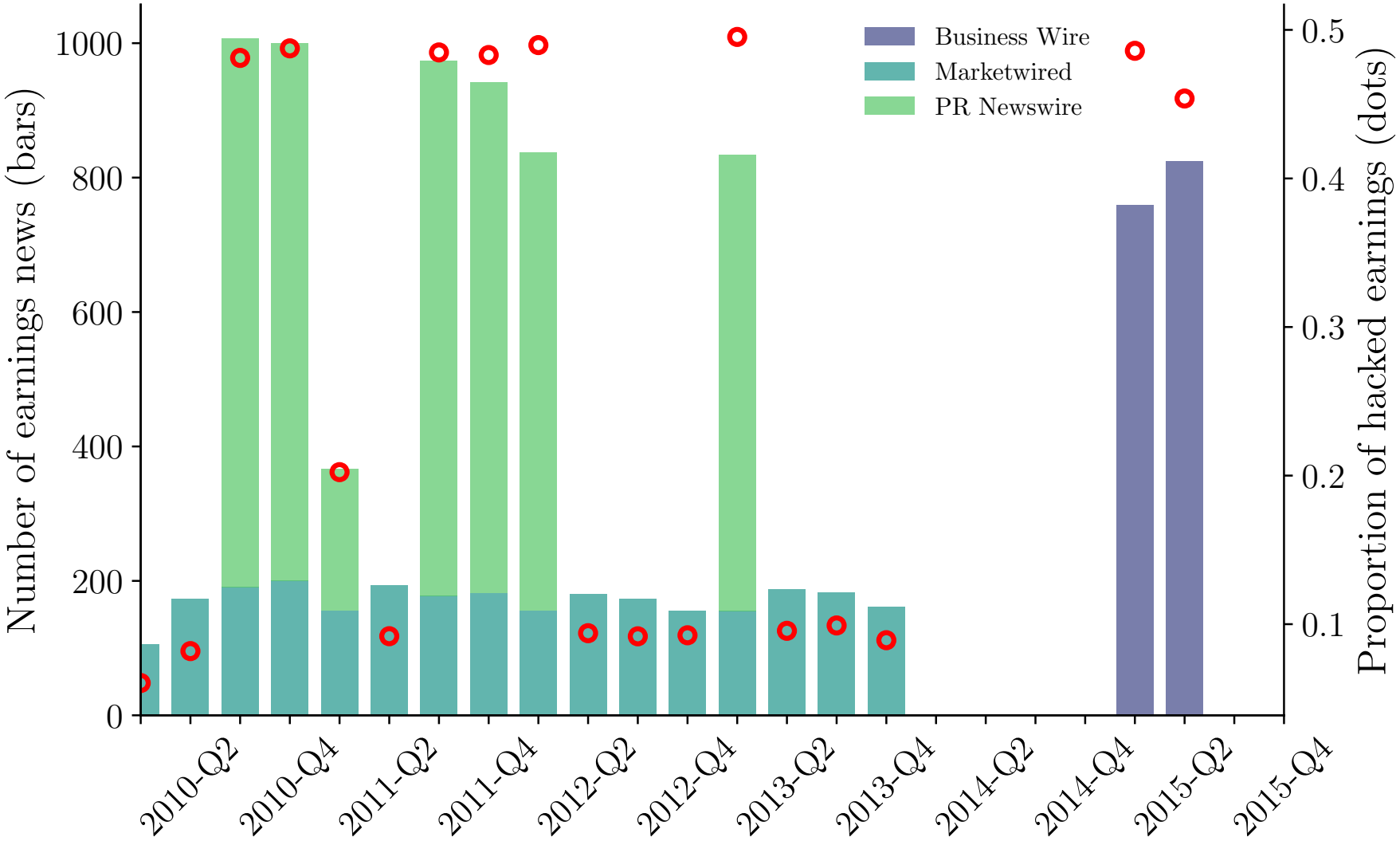
After-hours market



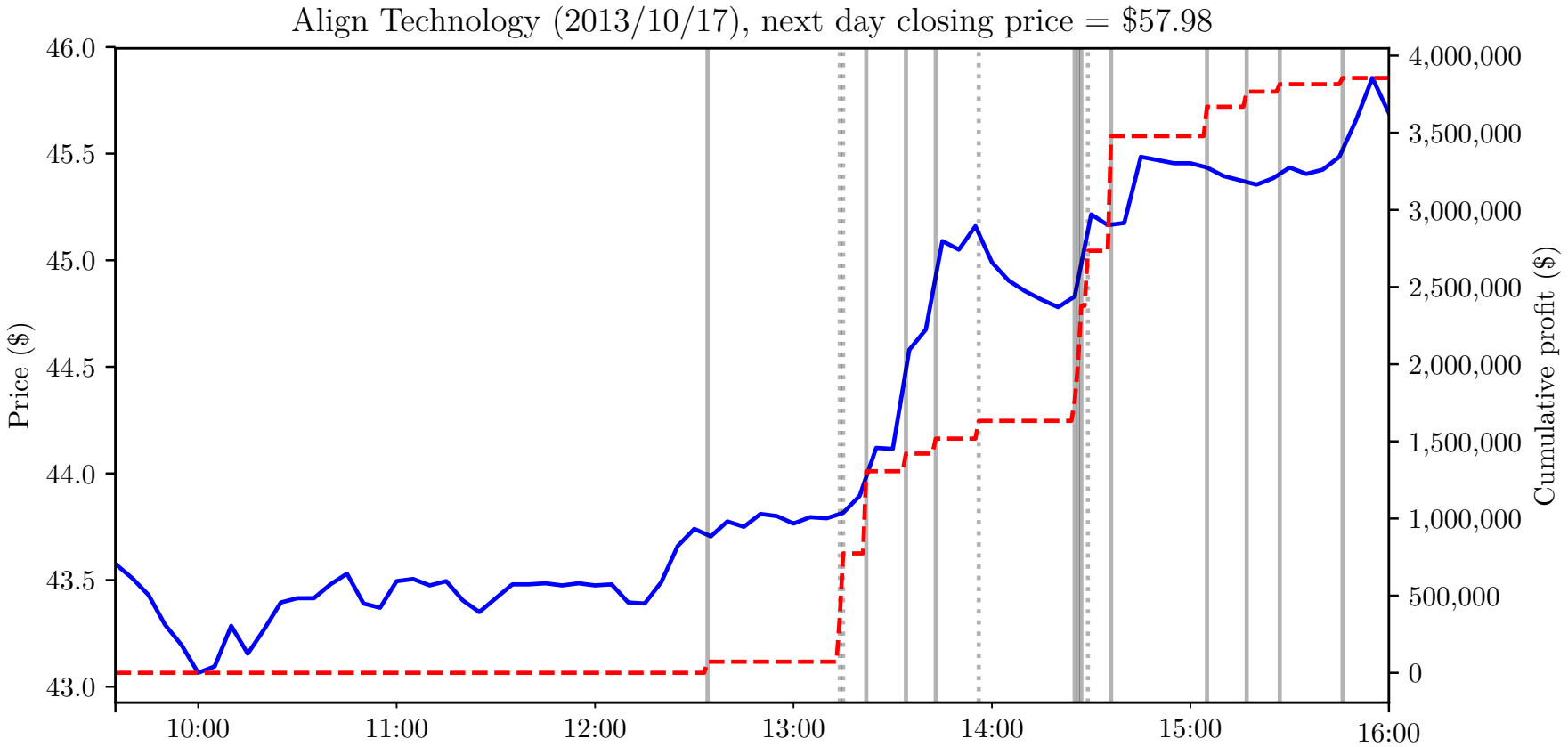
Data

1. Extract all earnings announcements in IBES from 2010 to 2015 for stocks in CRSP.
2. Assign to each earnings announcement the corresponding newswire company with Ravenpack with substantial validation work.
   - A total of $\sim$ 44,000 earnings announcements.
3. Retrieve intraday data from TAQ.
4. SEC documentation and legal filings to retrieve press releases and when were newswire companies exposed to hacks, expert witness reports, etc. ($>$4500 pages of court documents from PACER).

Number of earnings news exposed to hacks



Trading anecdotes

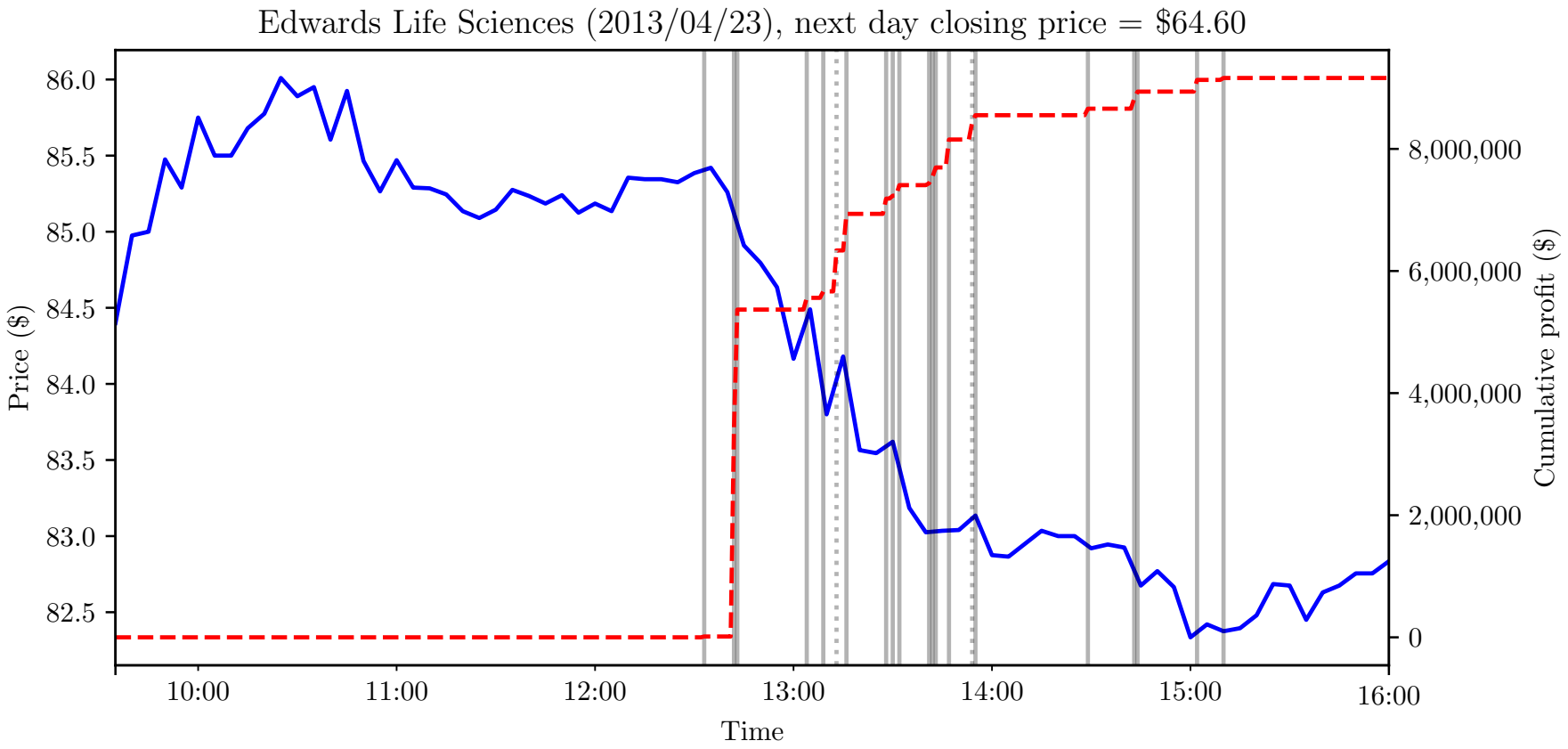


Stock price

Cumulative profits

Trades in stocks (solid) and derivative instruments (dotted)

Trading anecdotes



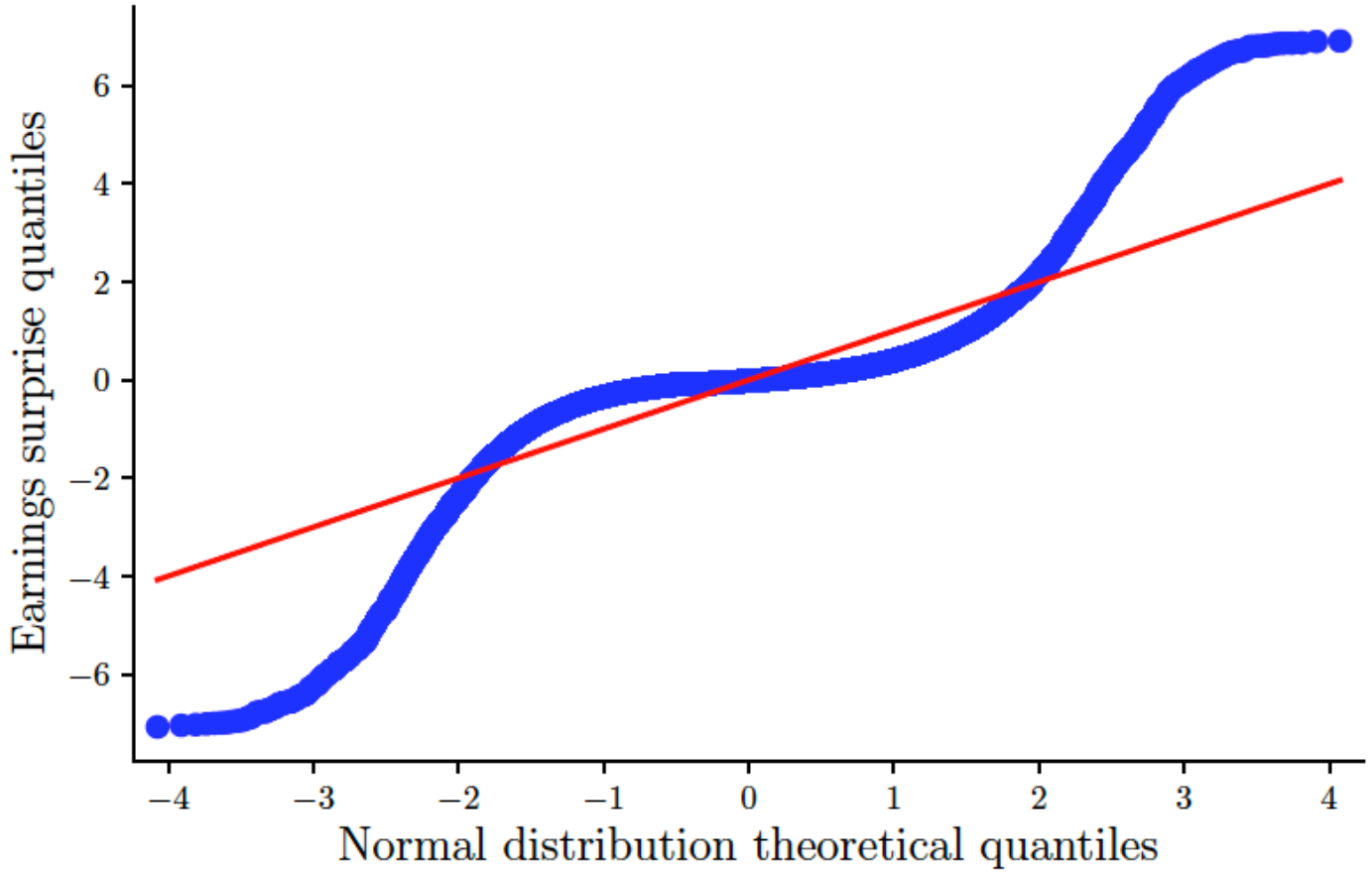
Stock price

Cumulative profits

Trades in stocks (solid) and derivative instruments (dotted)

Standardized Unexpected Earnings

$$Surprise_{i,t} = \frac{EPS_{i,t} - E_{t-1}[EPS_{i,t}]}{P_{i,t-5}},$$



Stock Selection

- Typical insiders only have access to information for one firm, and they decide to trade or not. We only observe those that trade and get caught.
- Our setting is unique:
  - We know the firms for which they had information, even if they did not trade on it.
  - We know the information they had access to (press releases).
  - We know *some* of the announcements they traded on (724).
- We use this to study how the hackers selected events to trade on.

Are firm characteristics correlated with hack periods?

$$Characteristics_{i,t} = \mathbf{1}_{[\text{Hacked}],i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t},$$

Panel B. With year-quarter fixed-effects

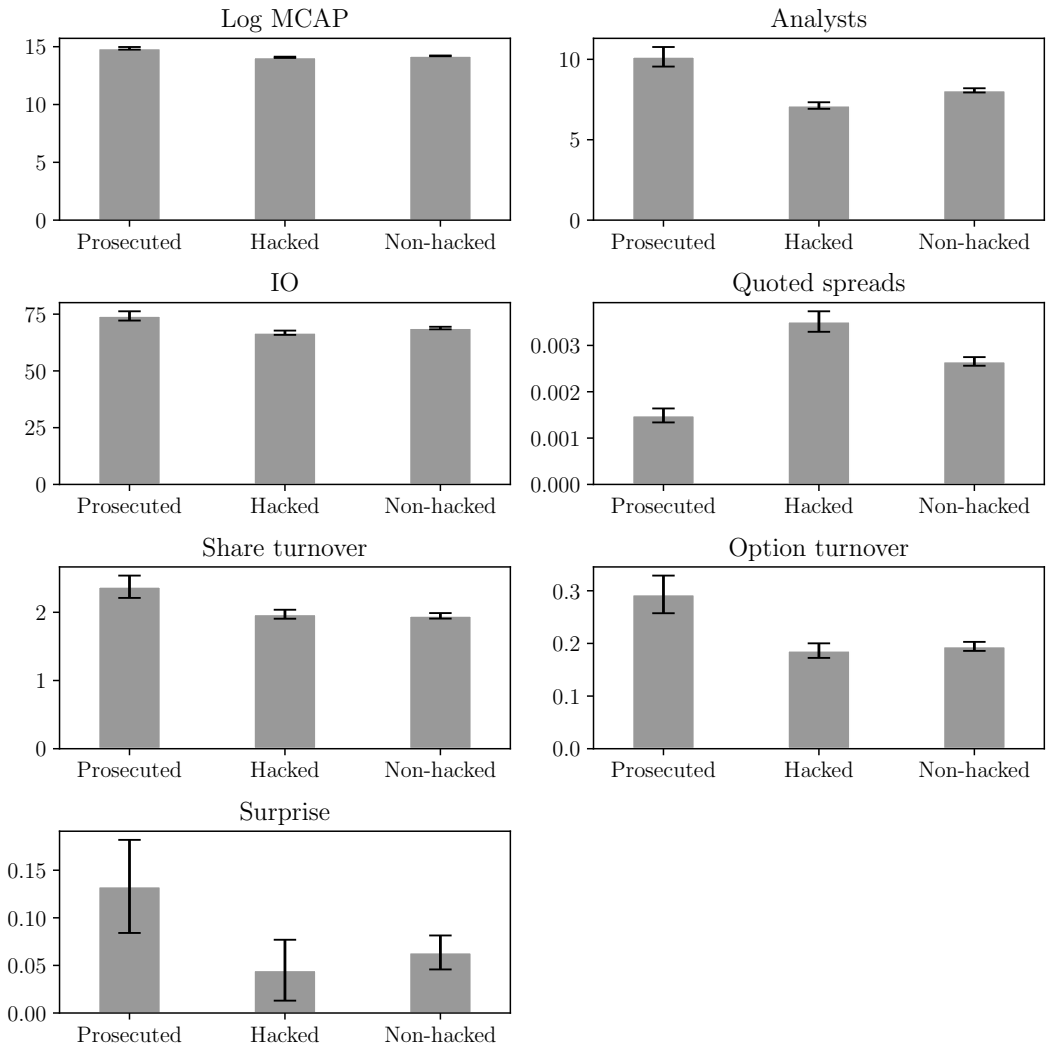
	Surprise (1)	|Surprise| (2)	Ln MCAP (3)	IO (4)	N. analysts (5)	Ln Q-value (6)	Share turn. (7)	Option turn. (8)
$\mathbf{1}_{[\text{Hacked}]}$	-0.000	0.000	-0.195**	-3.119**	-0.153	0.045	0.051	0.003
	(0.000)	(0.001)	(0.079)	(1.336)	(0.221)	(0.055)	(0.061)	(0.014)
N	43,687	43,687	43,687	43,687	43,687	35,273	43,687	43,687
Adjusted R^2	0.000	0.000	0.002	0.001	0.000	0.000	0.000	0.000
Year-Quarter F.E.	Y	Y	Y	Y	Y	Y	Y	Y
Firm F.E.	N	N	N	N	N	N	N	N

Panel C. With firm and year-quarter fixed-effects

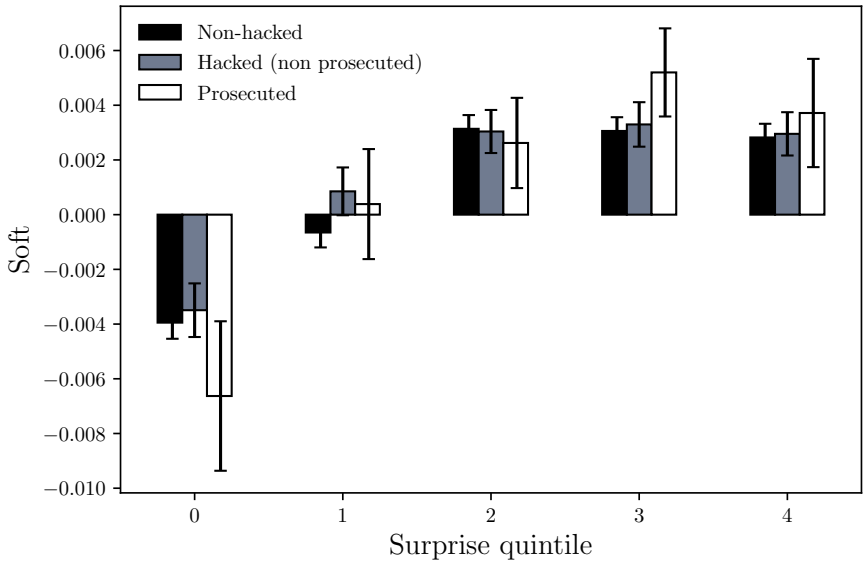
	Surprise (1)	|Surprise| (2)	Ln MCAP (3)	IO (4)	N. analysts (5)	Ln Q-value (6)	Share turn. (7)	Option turn. (8)
$\mathbf{1}_{[\text{Hacked}]}$	0.000	-0.000	0.002	0.084	0.088	-0.021	-0.010	-0.004
	(0.000)	(0.001)	(0.011)	(0.293)	(0.078)	(0.019)	(0.032)	(0.006)
N	43,687	43,687	43,687	43,687	43,687	35,273	43,687	43,687
Adjusted R^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Year-Quarter F.E.	Y	Y	Y	Y	Y	Y	Y	Y
Firm F.E.	Y	Y	Y	Y	Y	Y	Y	Y

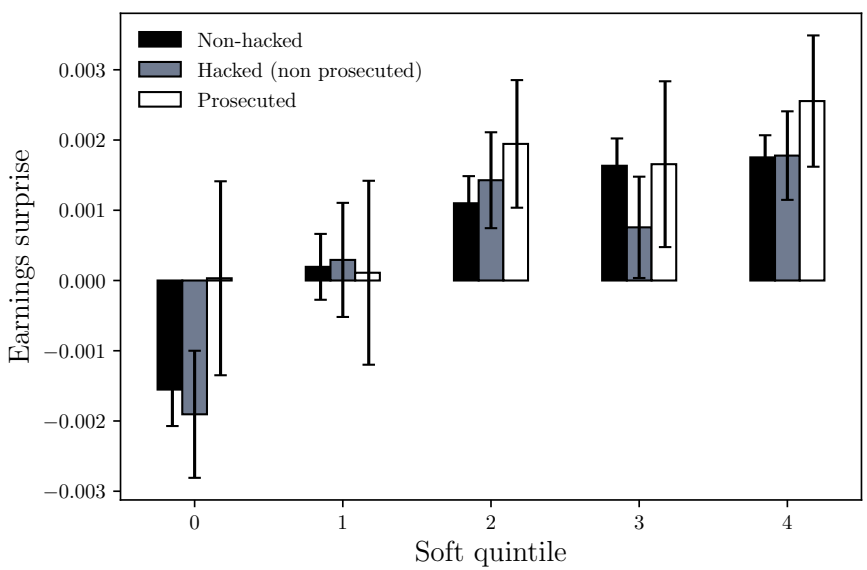
$\Rightarrow$ Hackers most likely did not select newswires because of differences in the trading environment of their client firms.

Are firm characteristics correlated with known cases?

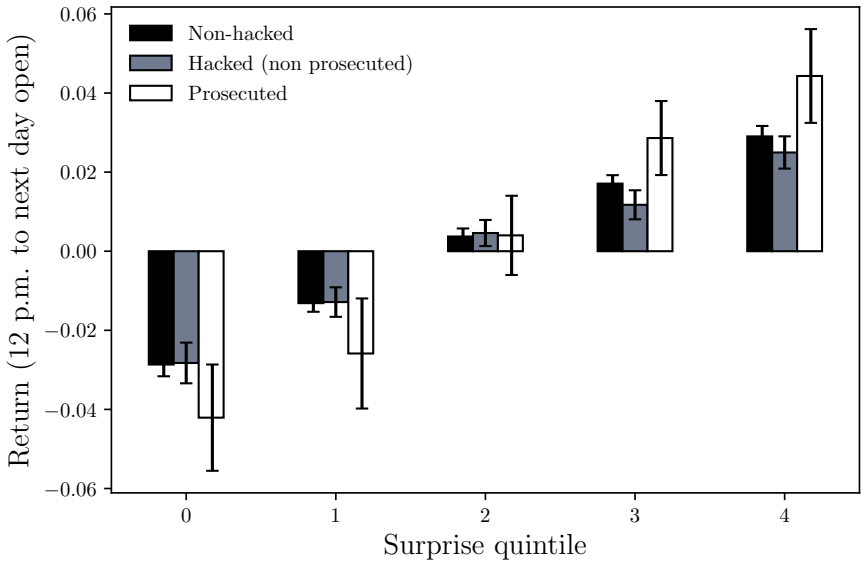


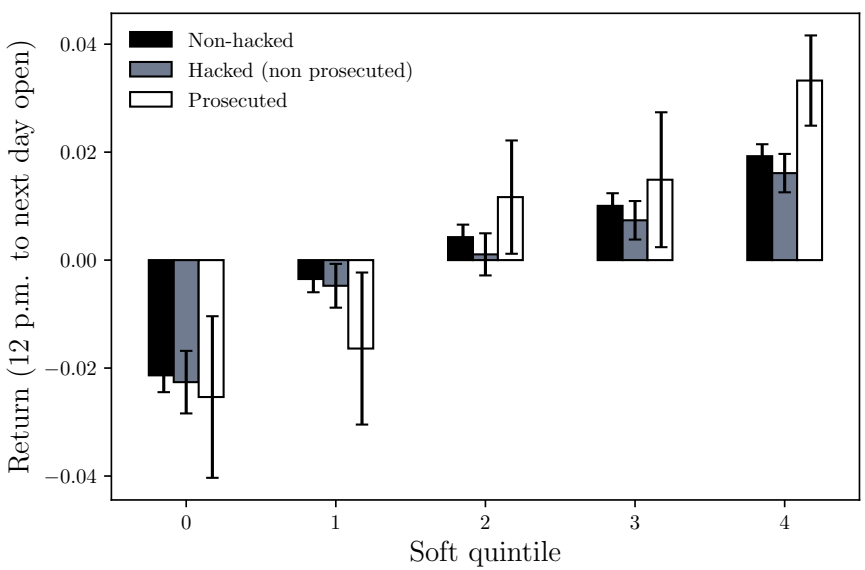
Is the information the same?





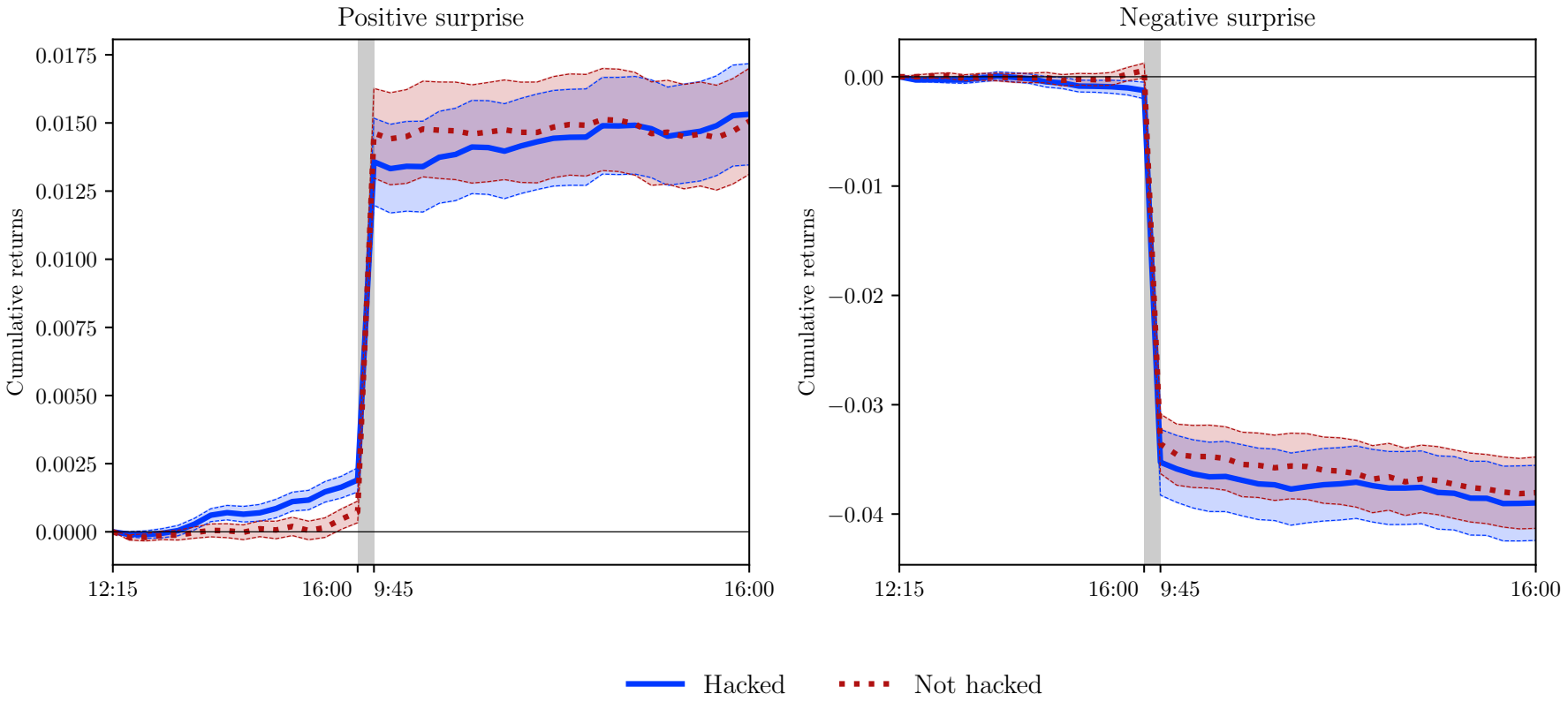
Are the returns the same?





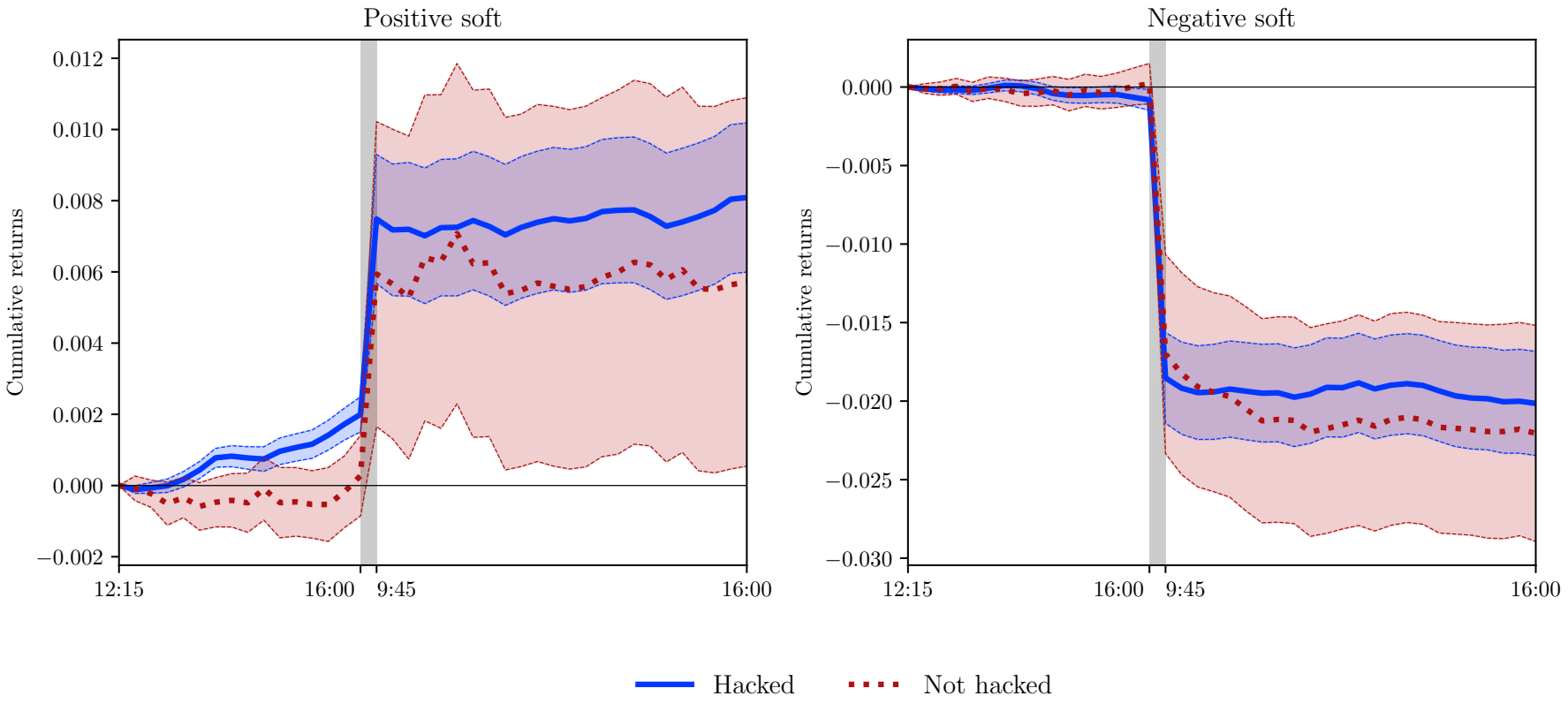
Price drifts before earnings announcements (S&P 1500)

Panel A. Earnings surprise



Price drifts before earnings announcements (S&P 1500)

Panel B. Soft information



Afterhours returns and informed trading

$$Return^{on}_{i,t} = \beta_1 Surprise_{i,t} + \beta_2 Surprise_{i,t} \times \mathbf{1}_{[\text{Hacked}]i,t} + \beta_3 \mathbf{1}_{[\text{Hacked}]i,t} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t}$$

	(1)	(2)	(3)	(4)
$Surprise$	1.364***	1.436***	1.443***	1.426***
	(0.073)	(0.079)	(0.079)	(0.062)
$Surprise \times \mathbf{1}_{[\text{Hacked}]}$	-0.212**	-0.234**	-0.241**	-0.215**
	(0.103)	(0.103)	(0.104)	(0.101)
$\mathbf{1}_{[\text{Hacked}]}$	-0.001	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
N	43,687	43,687	43,687	43,687
Adjusted R^2	0.062	0.060	0.071	0.067
Controls	N	N	Y	Y
Year-Quarter F.E.	Y	Y	Y	N
Firm F.E.	N	Y	Y	Y
Date F.E.	N	N	N	Y

- After-hour returns of stocks exposed to hacks are 15% less sensitive to earnings surprises.

Afterhours returns and informed trading

$$Return^{on}_{i,t} = \beta_1 Soft_{i,t} + \beta_2 Soft_{i,t} \times \mathbf{1}_{[\text{Hacked}]i,t} + \beta_3 \mathbf{1}_{[\text{Hacked}]i,t} + \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t}$$

	(1)	(2)	(3)	(4)
$Soft$	1.246***	1.352***	1.345***	1.334***
	(0.062)	(0.065)	(0.063)	(0.045)
$Soft \times \mathbf{1}_{[\text{Hacked}]}$	-0.184*	-0.220**	-0.222**	-0.210**
	(0.097)	(0.099)	(0.100)	(0.083)
$\mathbf{1}_{[\text{Hacked}]}$	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
N	36,750	36,750	36,750	36,750
Adjusted R^2	0.069	0.062	0.072	0.068
Controls	N	N	Y	Y
Year-Quarter F.E.	Y	Y	Y	N
Firm F.E.	N	Y	Y	Y
Date F.E.	N	N	N	Y

- After-hour returns of stocks exposed to hacks are 15% less sensitive to soft information.

Afterhours returns and informed trading

$$\begin{aligned} Return^{on}_{i,t} =& \beta_1 Surprise_{i,t} + \beta_2 Soft_{i,t} + \beta_3 Surprise_{i,t} \times \mathbf{1}_{[\text{Hacked}]i,t} + \\ & \beta_4 Soft_{i,t} \times \mathbf{1}_{[\text{Hacked}]i,t} + \beta_5 \mathbf{1}_{[\text{Hacked}]i,t} + \\ & \Gamma' Controls_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t} \end{aligned}$$

	(1)	(2)	(3)	(4)
$Surprise$	1.225***	1.326***	1.336***	1.308***
	(0.071)	(0.081)	(0.081)	(0.064)
$Soft$	1.090***	1.192***	1.183***	1.176***
	(0.060)	(0.060)	(0.059)	(0.042)
$Surprise \times \mathbf{1}_{[\text{Hacked}]}$	-0.174	-0.199*	-0.207*	-0.144
	(0.111)	(0.111)	(0.112)	(0.113)
$Soft \times \mathbf{1}_{[\text{Hacked}]}$	-0.179**	-0.202**	-0.202**	-0.193**
	(0.088)	(0.091)	(0.092)	(0.080)
$\mathbf{1}_{[\text{Hacked}]}$	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
N	36,750	36,750	36,750	36,750
Adjusted R^2	0.115	0.110	0.120	0.116
Controls	N	N	Y	Y
Year-Quarter F.E.	Y	Y	Y	N
Firm F.E.	N	Y	Y	Y
Date F.E.	N	N	N	Y

Did liquidity providers respond?

Classical models of microstructure: market makers observe order flow and informed traders respond by splitting trades to avoid punishment

- Challenging empirically because of strategic game
  - "Tests" of Kyle-style[1] behavior tend to focus on the *trader side*
- One needs to impose a constraint on trade splitting to test whether market makers can respond
  - If liquidity providers can charge higher spreads when they believe trade volume is informed, it is rational to trade split

1. We examine whether measures of order flow and spreads were higher when the hackers were trading
   - Order flow: Equity volume, option volume, equity turnover, absolute order imbalance
   - Spreads: Effective spreads, realized spreads, price impact, quoted spreads
2. We use two-stage least squares estimation to further examine how the *informedness* of order flow matters in determining spreads (not in paper)

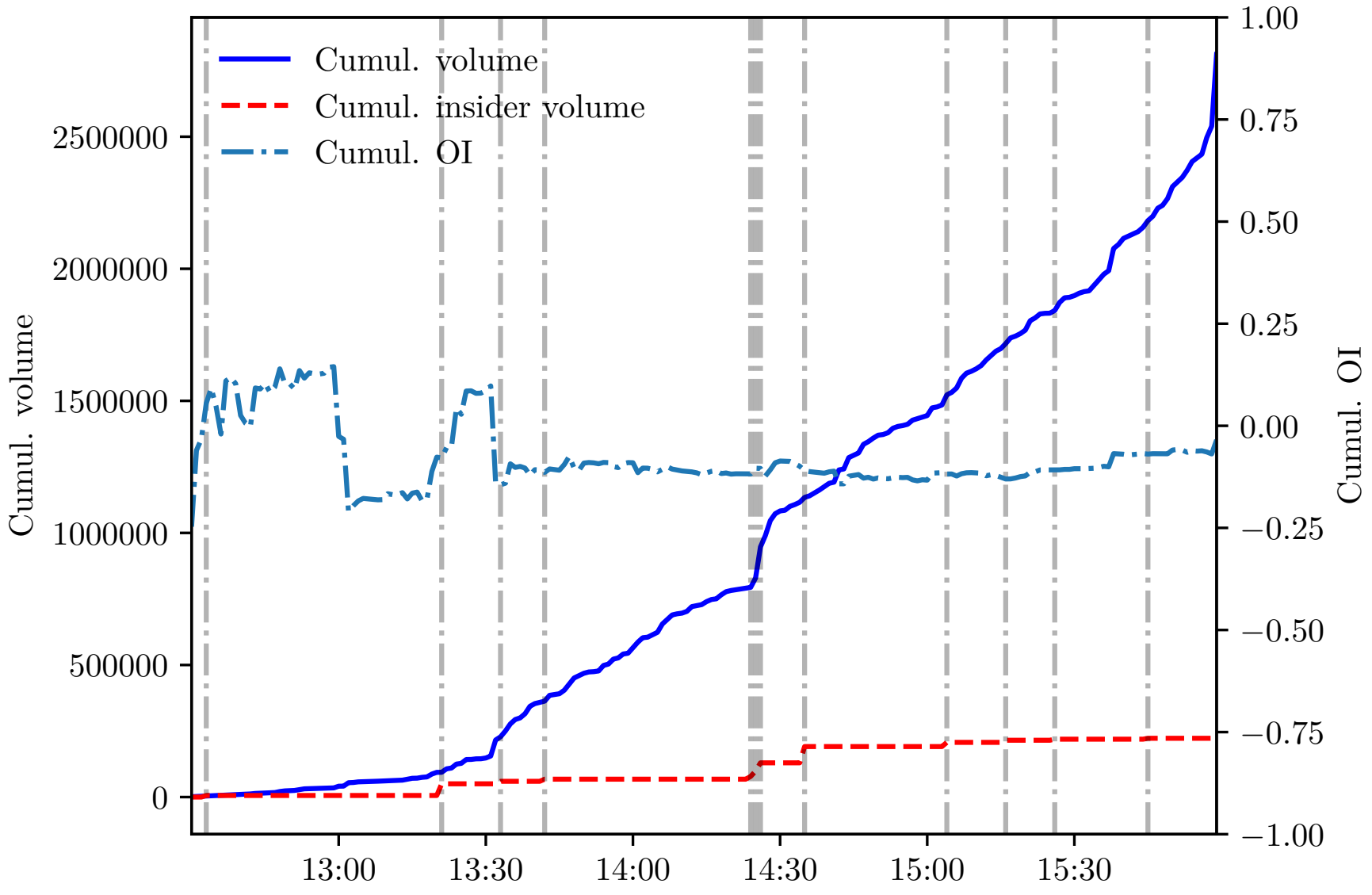
[1] Refers to the following classic paper: Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society, 1315-1335.

Order flow and spreads

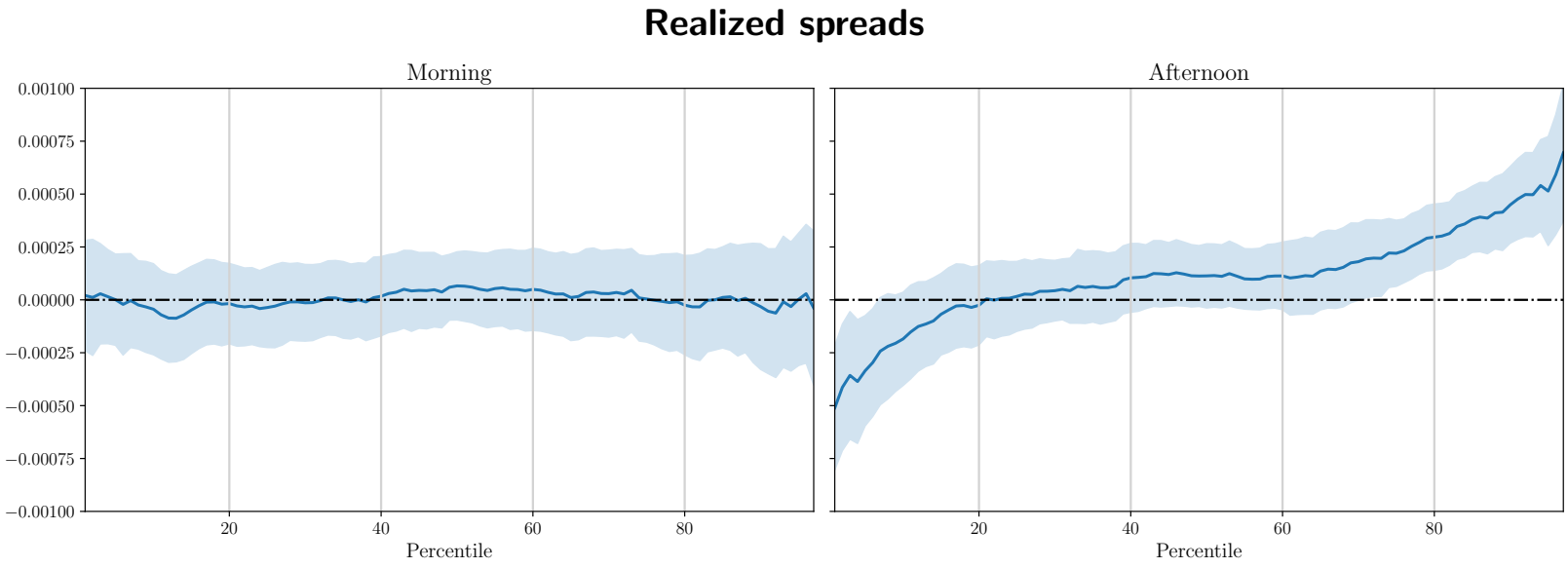
	Order flow measures				Spread measures			
	Turnover	Log(volume)	|OI|	Log(option vol)	Effective spread	Realized spread	Price impact	Quoted spread
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbf{1}_{[\text{Hacked}]}$	0.0490***	0.0350**	0.0080	0.0721**	0.0312***	0.0445***	0.0153	-0.0140
	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.02)
N	43,687	43,687	43,687	43,687	43,687	43,687	43,687	43,687
R^2	0.021	0.038	0.013	0.054	0.153	0.034	0.128	0.022
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Year-Quarter F.E.	Y	Y	Y	Y	Y	Y	Y	Y
Firm F.E.	Y	Y	Y	Y	Y	Y	Y	Y

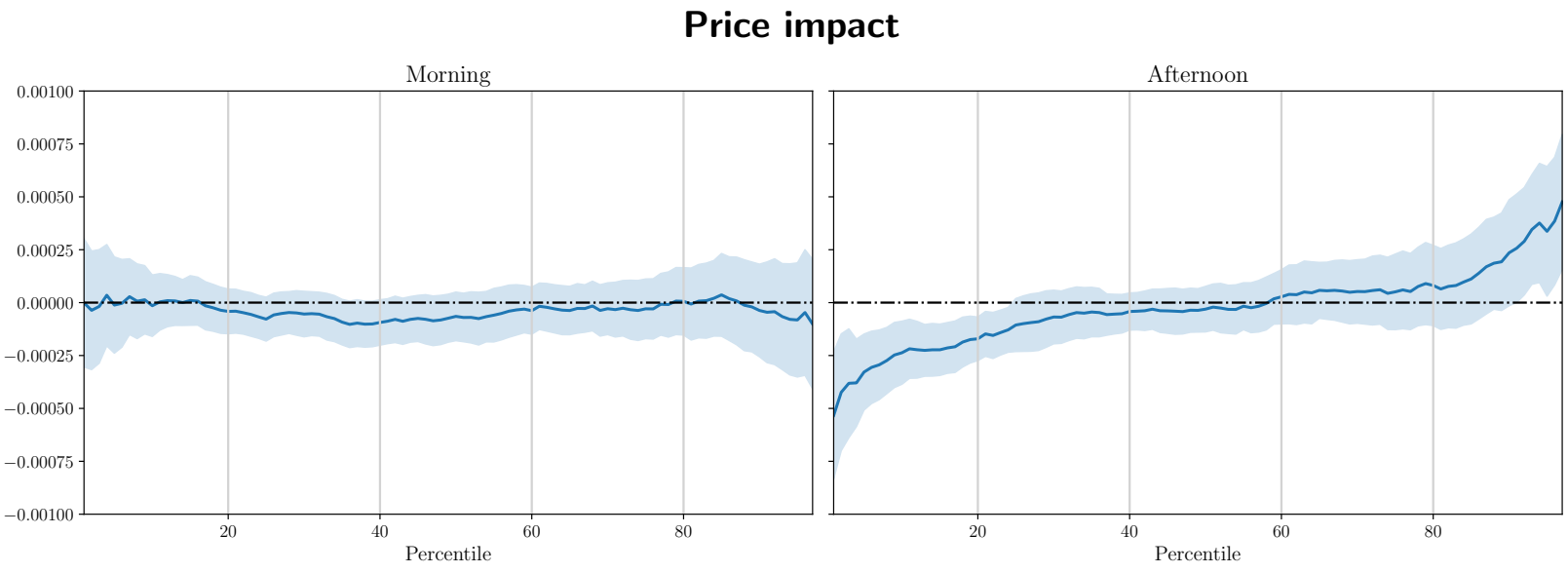
- Most order flow measures robustly higher in the afternoon (when the hackers were most often active)
  - Equity volume, option volume, equity turnover are higher
  - Absolute order imbalance not higher
    - Possibly due to inventory balancing
- Effective and realized spreads also higher; price impact not
  - Consistent with liquidity providers charging liquidity takers more
- No differential patterns in the morning (before the hackers typically had access to news)

ALIGN Volume and Order Imbalance



How does the spread distribution change?





- Realized spreads increase from about the **70th percentile**

The informedness of order flow and spreads

- Previous results show that several measures of order flow and spreads were higher when informed hackers were trading. Implies a positive relationship between order flow and spreads
  - Consistent with market making behavior in classic models when informed traders face constraints on trade splitting
- However, the correlation between order flow and spreads is likely to represent several different associations
  - Order flow is likely to be higher in stocks that are easier to trade (-)
  - Noise trading likely to be more prevalent with lower spreads (-)
  - Market makers are likely to to charge higher spreads when they face more informed trading (+)
- We examine the correlation between order flow and spreads using this setting to isolate *informed* order flow
  - Isolate *informed* variation in order flow in a 2SLS framework and compare it to the *unconditional* OLS correlation between order flow and spreads

OLS relationship between order flow and spreads

		Dependant variable		
	Regression indep. variable	Effective spreads (1)	Price impact (2)	Realized spreads (3)
(I)	Turnover	-0.0304***	0.0054	-0.0388***
		(0.0062)	(0.0055)	(0.0065)
(II)	Log(volume)	-0.1513***	0.0151	-0.1713***
		(0.0129)	(0.0134)	(0.0137)
(III)	Log(option vol)	-0.0057***	-0.0076**	-0.0023
		(0.0015)	(0.0031)	(0.0021)

- Unconditionally order flow is negatively associated with spreads across all measures

2SLS estimation

First Stage:

$$Order\,Flow_{i,t} = \beta \mathbf{1}_{[\text{Hacked}]i,t} + \Gamma' Controls_{i,t} + \alpha_t + \alpha_i + \varepsilon_{i,t},$$

Second Stage:

$$Spreads_{i,t} = \delta \widehat{Order\,Flow}_{i,t} + \Gamma' Controls_{i,t} + \alpha_t + \alpha_i + \varepsilon_{i,t}.$$

- 2SLS procedure *only* uses varation in order flow measures due to the hacking scandal ($\mathbf{1}_{[\text{Hacked}]i,t}$)
  - Has the effect of isolating (one type of) *informed order flow*

2SLS second stage results

		Dependant variable		
	Regression indep. variable	Effective spreads (1)	Price impact (2)	Realized spreads (3)
(I)	Turnover	0.5220**	-0.3721	0.8338**
		(0.2432)	(0.3920)	(0.3624)
(II)	Log(volume)	0.8444*	-0.6015	1.3460*
		(0.5125)	(0.6862)	(0.7676)
(III)	Log(option vol)	0.3380**	-0.2410	0.5407**
		(0.1601)	(0.2738)	(0.2694)

- First stage results are the association between $\mathbf{1}_{[\text{Hacked}]i,t}$ and order flow shown earlier
- Increased *informed* order flow is associated with higher spreads
  - More direct evidence that the market makers respond to increased informed order flow by charging higher speads ‘

Conclusion

Cyber risks expose financial markets to systematic information leakage and insider trading.

Our analysis shows that during the newswire hacking scheme:

1. Insiders chose to trade stocks with large surprises, those with better information environments and liquidity.
2. Price discovery for stocks exposed to hacks occurred before the earnings announcements.
3. Stock prices exposed to hacks were 15% less responsive to earnings surprises (as high as 50% for events reported by the SEC).

Several measures of order flow, effective and realized spreads were higher

- Provides evidence that liquidity providers were able to detect an increase in informed order flow and respond as several theories would predict