Talking to Digital Twins

Selective Disclosure and Belief Measurement in Financial Social Media

Boone Bowles

Raymond Duch

Sorin Sorescu







Brigham Young University
November 13, 2026

Talking to Digital Twins

Selective Disclosure and Belief Measurement in Financial Social Media

Boone Bowles

Raymond Duch

Sorin Sorescu

TEXAS A&M UNIVERSITY
Mays Business School

TEXAS A&M UNIVERSITY
Mays Business School

Brigham Young University
November 13, 2026

What do we learn from silence?

Stock A



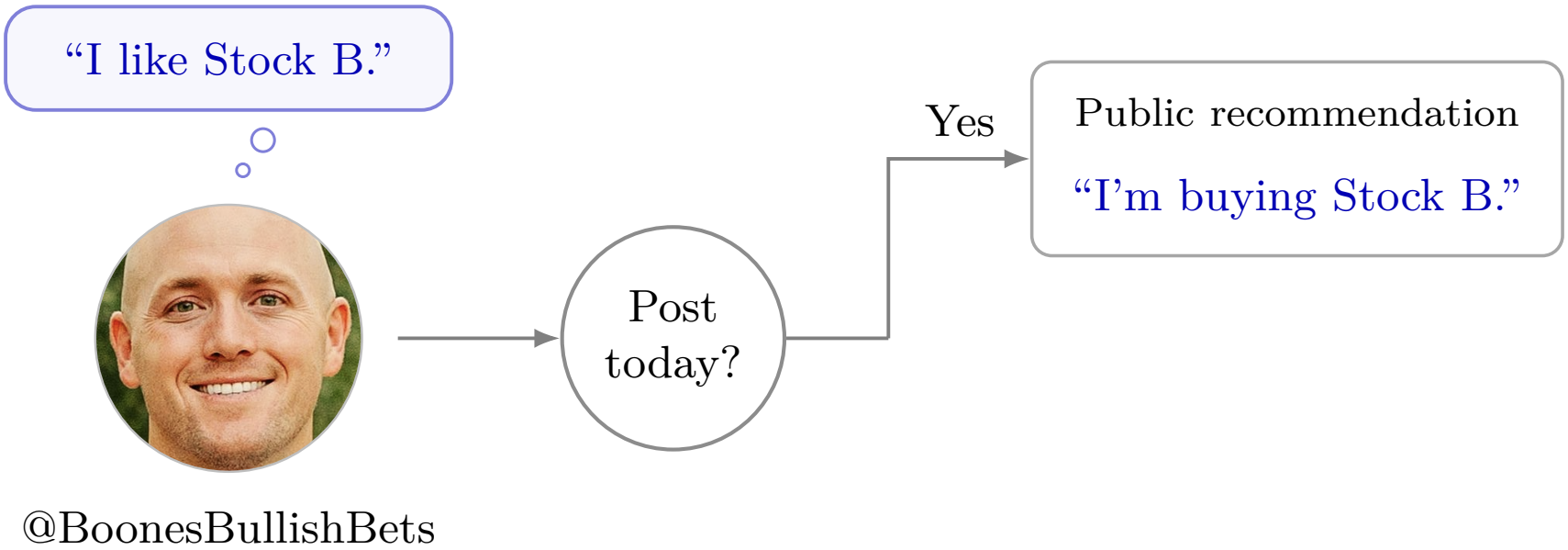
Probably bullish

Stock B

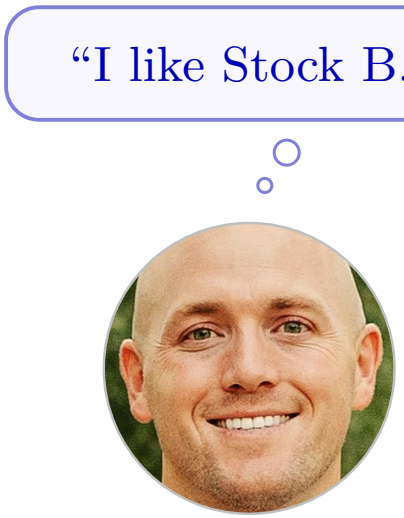


Bullish? Bearish? No opinion?
Not wanting to disclose?

What do we learn from silence?



What do we learn from silence?



@BoonesBullishBets

Post today?

Yes

Public recommendation

"I'm buying Stock B."

No

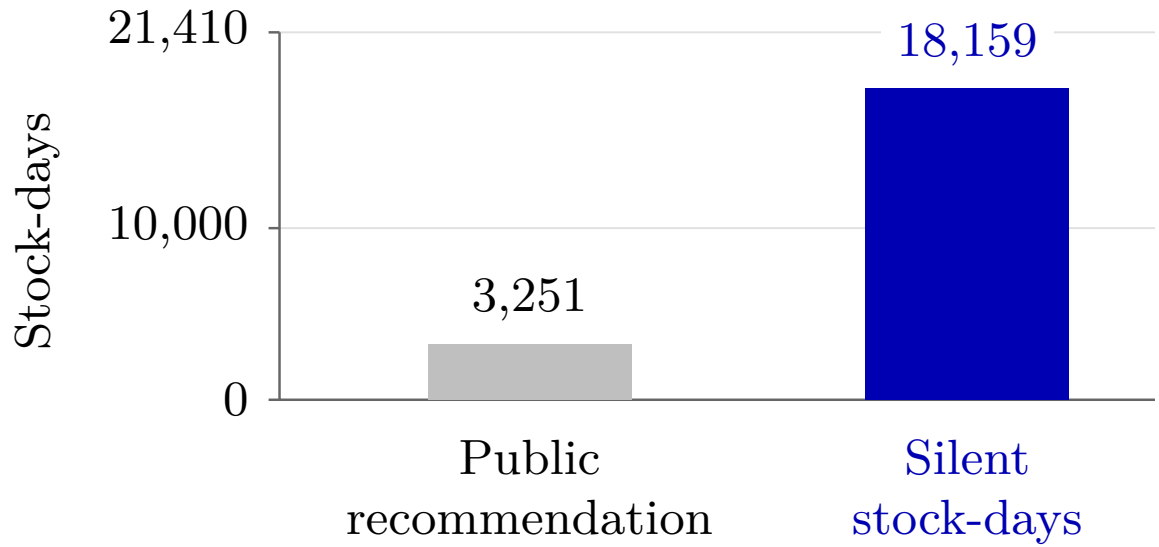
No post

- Audience won't care
- Not my niche
- Nothing new to add
- Not worth my time
- I'd rather trade myself

Silence does not reveal whether a view is absent or unspoken.

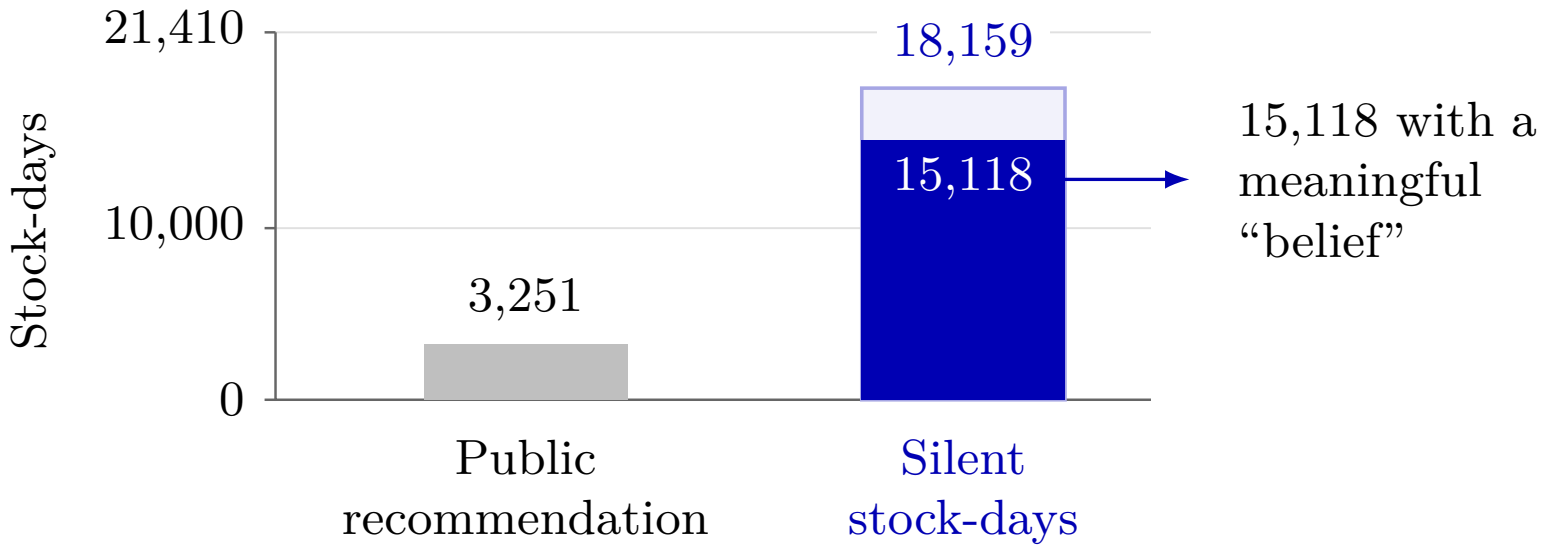
The silent region is large

Tracking posts from 81 X-finfluencers for 429 stocks and 50 trading days



The silent region is large

Tracking posts from 81 X-finfluencers for 429 stocks and 50 trading days



Can digital twin interviews elicit economically meaningful belief proxies in the silent region?

Research questions and main findings

1 Do interview responses reflect the public persona (the finfluencer)?

Responses align with public recommendations and retain account-specific patterns.

2 Are belief proxies recovered from the silent region informative?

More buy-leaning responses predict higher future returns.

3 What do disagreement and weak conviction predict?

Disagreement predicts lower returns and higher volatility; weak conviction predicts lower volatility.

Contribution to the literature

Literature	**This paper**
Financial social media Posts reveal disagreement and uneven skill. Cookson–Niessner (2020); Kakhbod et al. (2023)	Measure the silent region through daily interviews.
Beliefs and expectations Survey beliefs help explain investor behavior. Greenwood–Shleifer (2014); Giglio et al. (2021)	Daily, stock-level public-persona proxies.
LLMs and digital twins Models extract signals and simulate responses. Lopez-Lira–Tang (2024); Park et al. (2024)	Interview in real time; test future outcomes.

Measure the silent region; test its economic content.

Roadmap

1. **Interview design and data**
   How do we elicit and record responses?

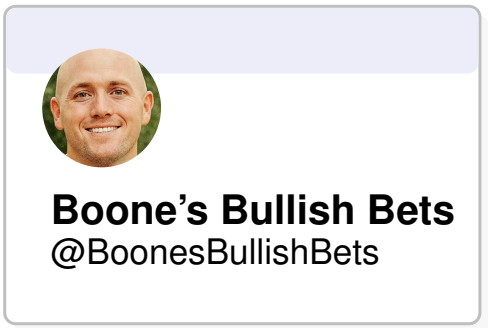
2. **Validation**
   Do responses reflect the public persona?

3. **Cross-sectional return predictability**
   Do responses predict future stock returns?

4. **Beyond direction**
   Disagreement, weak conviction, and macro evidence.

How do we make the silent region measurable?

Gather public inputs

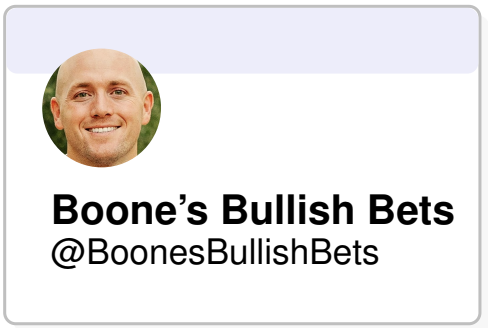


Profile data and recent posts.

- 81 accounts selected in June 2025
- $\geq$5,000 followers
- Directional posts on stocks and markets

How do we make the silent region measurable?

Gather public inputs



Profile data and recent posts.

- 81 accounts selected in June 2025
- $\geq$5,000 followers
- Directional posts on stocks and markets

Conduct LLM interview

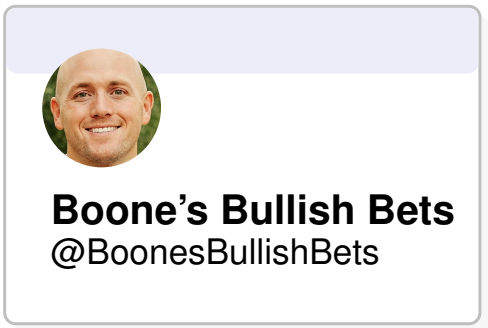


Digital twin instructed to answer questions as finfluencer.

- Macro questions
- 429 recurring large-cap stocks

How do we make the silent region measurable?

Gather public inputs



Profile data and recent posts.

- 81 accounts selected in June 2025
- $\geq$5,000 followers
- Directional posts on stocks and markets

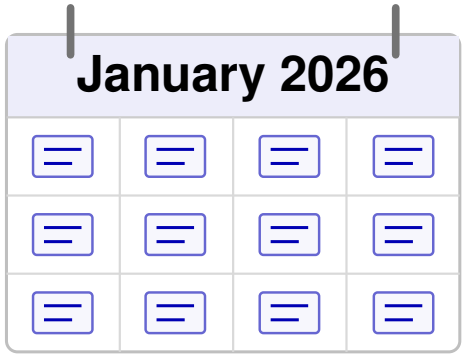
Conduct LLM interview



Digital twin instructed to answer questions as finfluencer.

- Macro questions
- 429 recurring large-cap stocks

Repeat daily



Every digital twin! Every day!

- 50 stock trading dates
- December 15, 2025–March 16, 2026

Example interview questions

Stock question

```
Based on your general knowledge of
market conditions, indicate whether you
would BUY, HOLD, or SELL (SHORT SELL)
the stock ticker:
[COMPANY NAME (TICKER)].
```



Example response

```
question:
[COMPANY NAME (TICKER)]

recommendation:  75

speculation:  30
```

Example interview questions

Macro questions: Q1 and Q3

Q1. What is your best estimate of the probability that the U.S. economy will enter a recession in the next 12 months?

Q3. In the next six months, do you expect business conditions to be better, worse, or the same?



Example response (Q1)

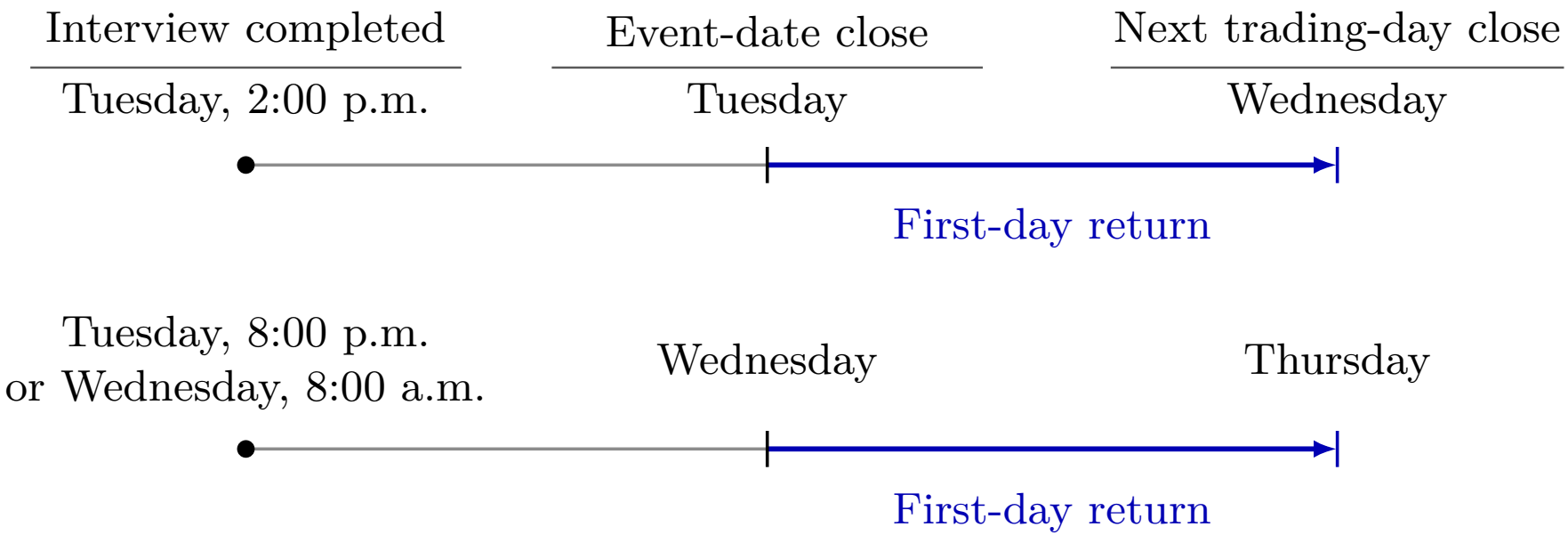
symbol: A2

category: 21 to 40% (Unlikely)

speculation: 30

Timing: interviews come before returns

Record responses first. Return windows later.



Roadmap

1. Interview design and data
   How do we elicit and record responses?

2. Validation
   Do responses reflect the public persona?

3. Cross-sectional return predictability
   Do responses predict future stock returns?

4. Beyond direction
   Disagreement, weak conviction, and macro evidence.

Do twin responses match public recommendations?

Same account · same stock · same date

Compare

Public buy/sell ⟷ Twin buy/sell

All non-neutral interviews
$N = 3{,}271$; score $\neq 50$

Meaningful, low-speculation
Score ≤ 35 or ≥ 65
Speculation ≤ 40; $N = 2{,}439$

Do twin responses match public recommendations?

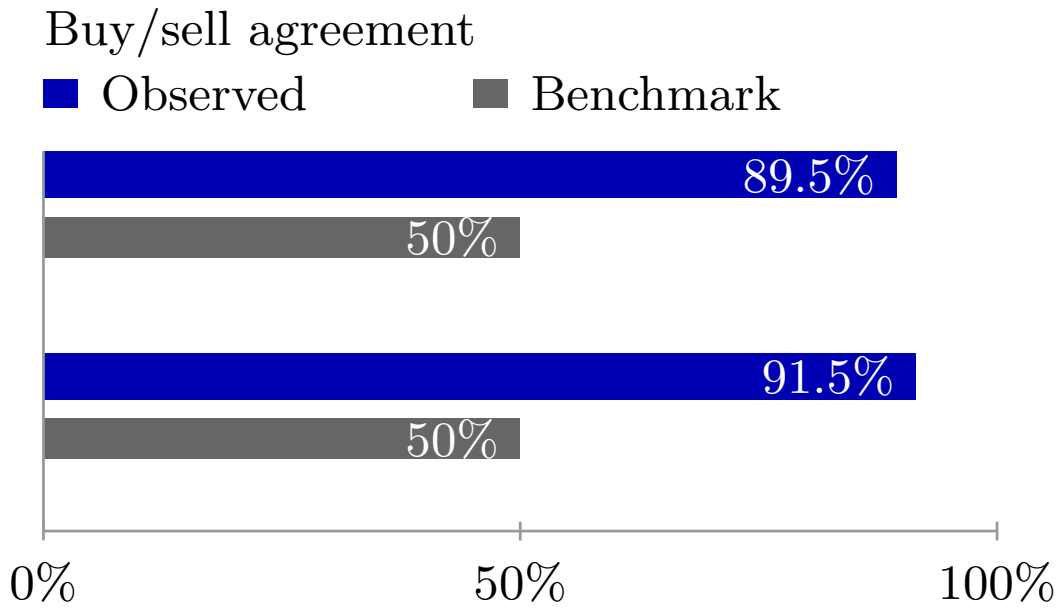
Same account · same stock · same date

Compare

Public buy/sell ⟷ Twin buy/sell

All non-neutral interviews
$N = 3{,}271$; score $\neq 50$

Meaningful, low-speculation
Score ≤ 35 or ≥ 65
Speculation ≤ 40; $N = 2{,}439$



Public recommendations and interview directions usually agree.

Are twin responses account-specific?

Seven macro questions: 0–100 bull scores (50 neutral; higher more bullish).

Subtract each question's daily mean across accounts.

Account A: residualized responses

	Date 1	Date 2	Date 3	Date 4
Q1	10	11	8	9
Q2	−6	−5	−4	−7
Q3	8	9	6	7
Q4	−4	−3	−2	−5
Q5	6	7	4	5
Q6	−2	−1	−4	−3
Q7	4	5	2	3

Are twin responses account-specific?

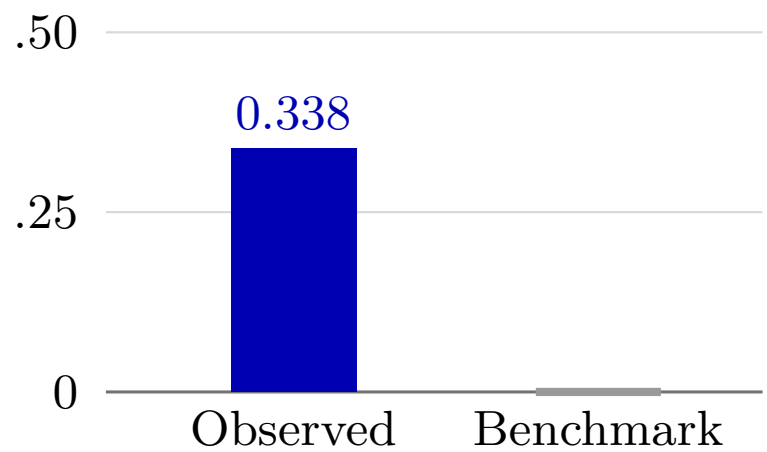
Seven macro questions: 0–100 bull scores (50 neutral; higher more bullish).

Subtract each question's daily mean across accounts.

Account A: residualized responses

	Date 1	Date 2	Date 3	Date 4
Q1	10	11	8	9
Q2	−6	−5	−4	−7
Q3	8	9	6	7
Q4	−4	−3	−2	−5
Q5	6	7	4	5
Q6	−2	−1	−4	−3
Q7	4	5	2	3

Increase in total-score R^2



$$\text{BullScore}_{iqt} = \gamma_{qt} + \alpha_i + \varepsilon_{iqt}$$

Account effects explain variation beyond common daily answers.

Are twin responses account-specific?

Seven macro questions: 0–100 bull scores (50 neutral; higher more bullish).

Subtract each question's daily mean across accounts.

Account A: residualized responses

	Date 1	Date 2	Date 3	Date 4
Q1	10	11	8	9
Q2	−6	−5	−4	−7
Q3	8	9	6	7
Q4	−4	−3	−2	−5
Q5	6	7	4	5
Q6	−2	−1	−4	−3
Q7	4	5	2	3

Training dates
Every other date.

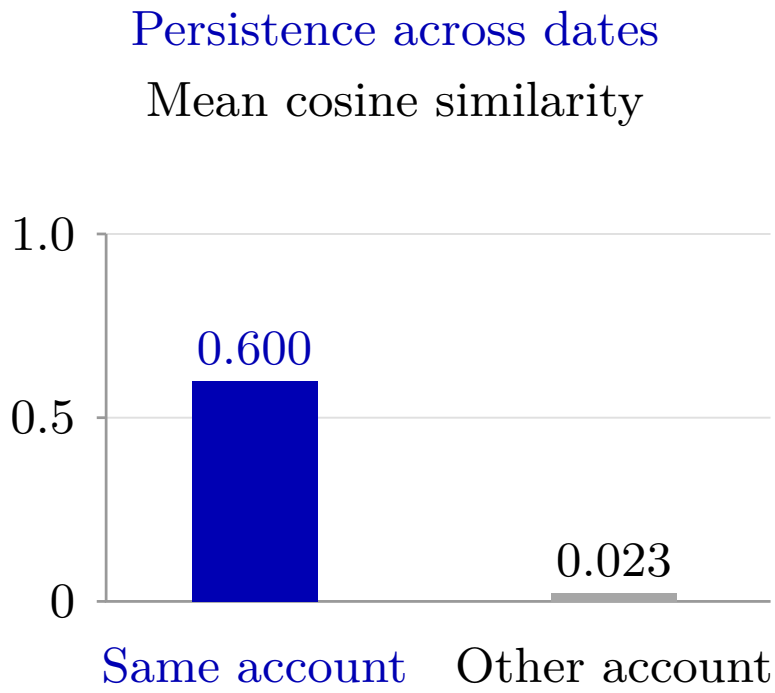
Held-out dates
Remaining dates.

For each account, average each question over training dates.
Match each held-out response vector to the most similar account profile.

$$\text{cosine}(x, y) = \frac{x^{\mathsf{T}} y}{\|x\| \, \|y\|}$$

x: held-out response vector
y: account's training profile

Account-specific patterns persist across interviews

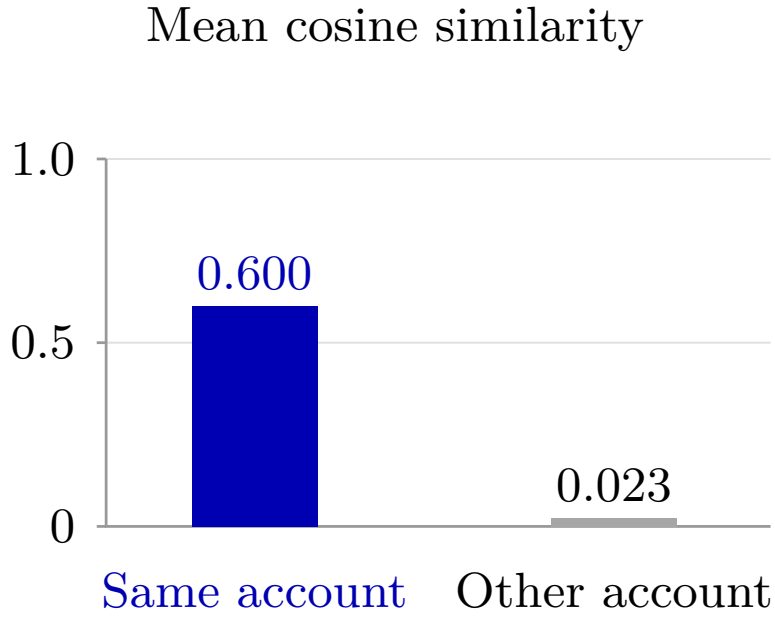


Compare on the next observed date:
same account versus another account.

The same account's response pattern persists across dates.

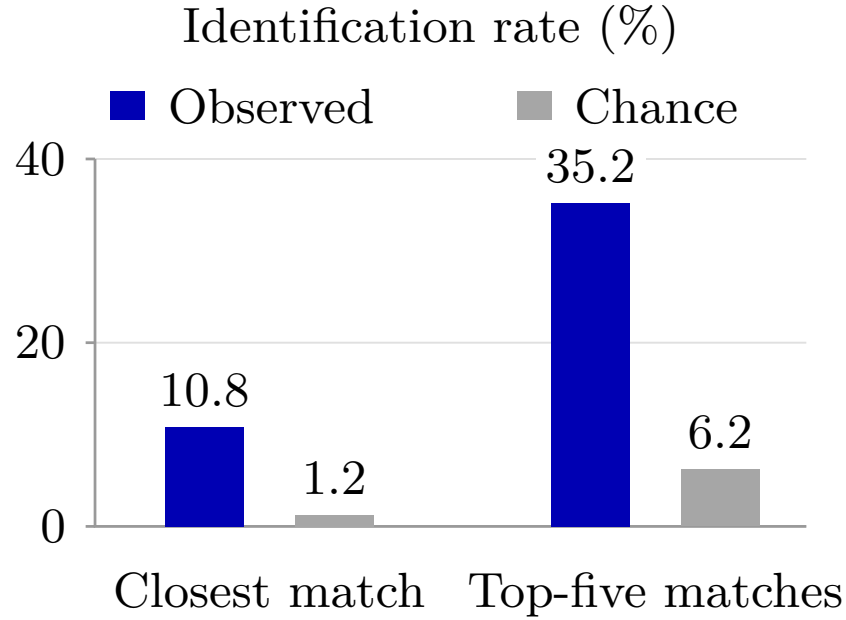
Account-specific patterns persist across interviews

Persistence across dates



Compare on the next observed date: same account versus another account.

Held-out account identification



Chance among 81 profiles: 1/81 closest; 5/81 top five.

Response patterns persist and help identify the account.

Better control samples

Could a generic model produce the same patterns?

	Regular interviews per day	
Respondent	No web	Web permitted
Bare model No finfluencer persona	25	25
Generic finfluencer A financial-influencer role	25	25
Digital twin Named persona, profile and posts	75	75

Additional experiment: October 7–November 1, 2026; preliminary.

Better control samples

Could a generic model produce the same patterns?

Respondent	Regular interviews per day: No web	Regular interviews per day: Web permitted
Bare model No finfluencer persona	25	25
Generic finfluencer A financial-influencer role	25	25
Digital twin Named persona, profile and posts	75	75

Additional experiment: October 7–November 1, 2026; preliminary.

Focus on digital twins and generic interviews with web permitted.

Can stock-specific answers identify the persona?

Use stock scores in the same held-out matching test.

Subtract each stock's daily mean across interviews.

Persona A: residualized stock scores

	Date 1	Date 2	Date 3	Date 4
Stock A	12	11	8	9
Stock B	−4	−5	−6	−5
Stock C	6	5	4	6
	⋮	⋮	⋮	⋮

Training dates
Every other date.

Held-out dates
Remaining dates.

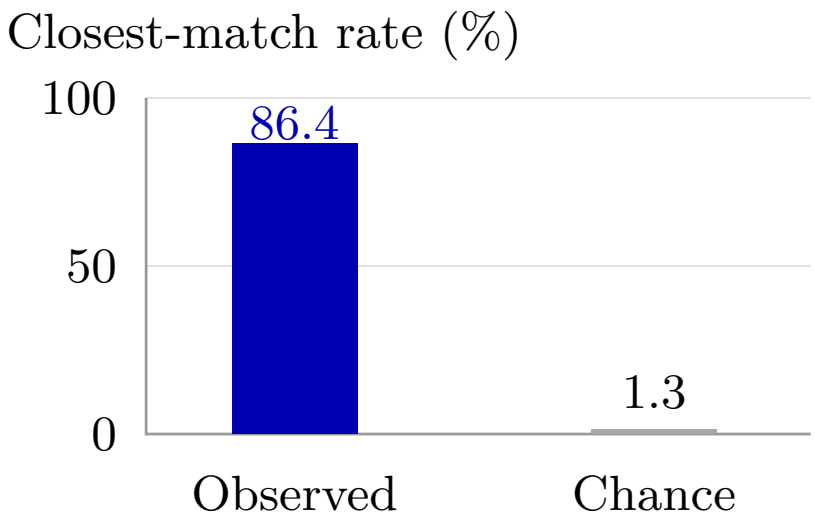
Average each stock over training dates to form 75 persona profiles.

Cosine similarity: does a held-out interview match the correct persona?

Repeat for generic interviews, using 25 arbitrary labels instead of personas.

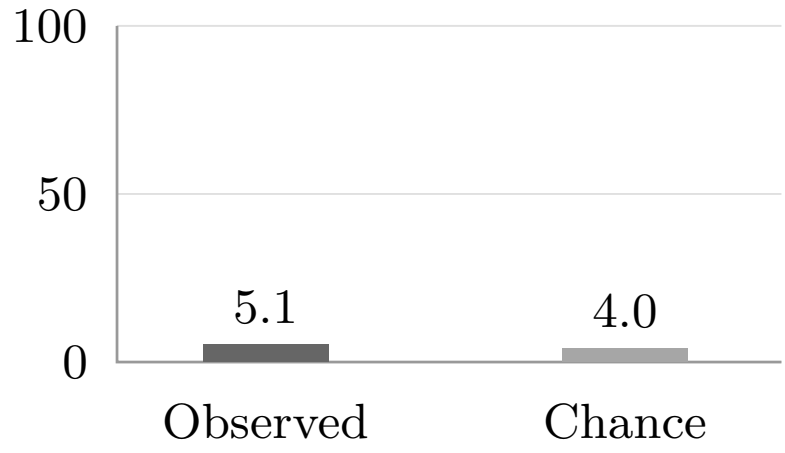
Can stock-specific answers identify the persona?

75 digital-twin personas



Chance among 75 personas:
1/75 for the closest match.

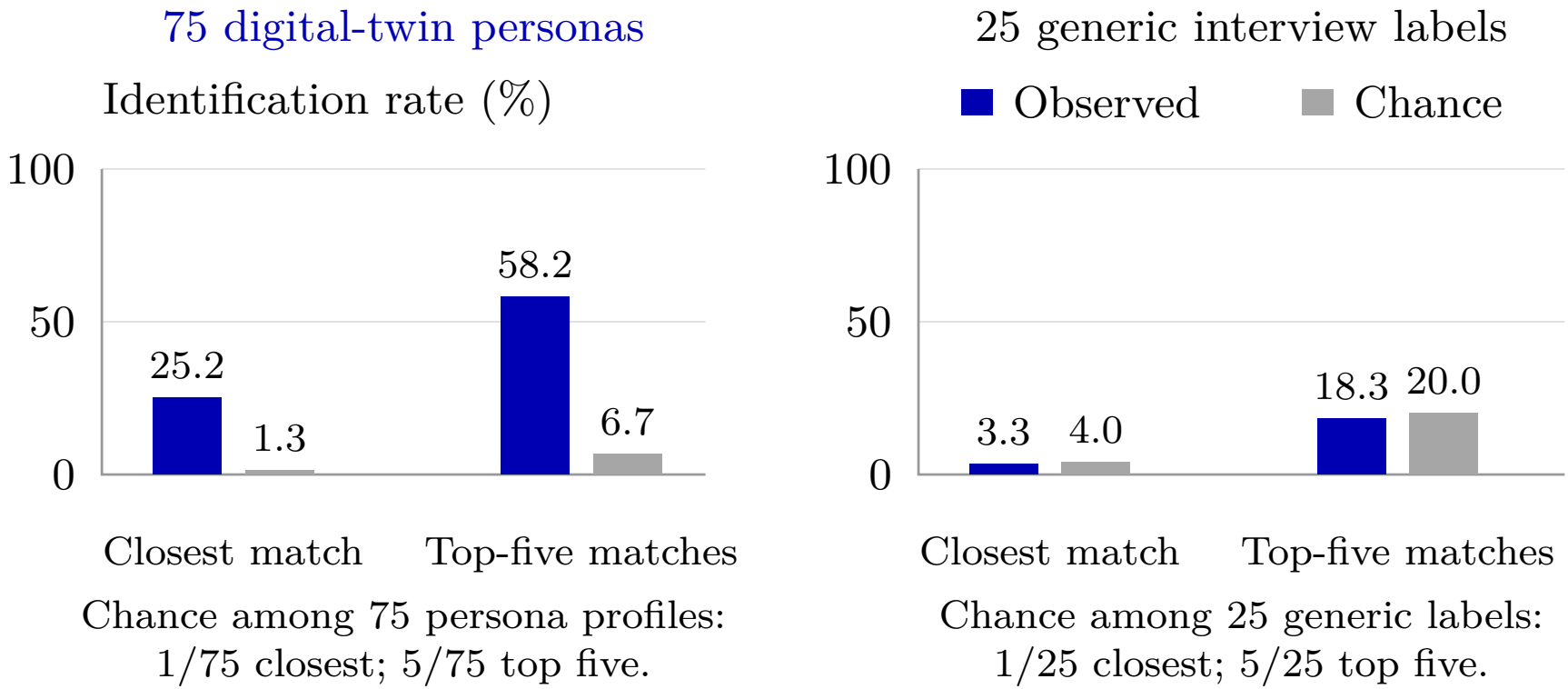
25 generic interview labels



Chance among 25 arbitrary labels:
1/25 for the closest match.

Twin answers retain distinctive patterns across dates.

Does the pattern extend to macro answers?



Held-out macro responses help identify the persona.

Roadmap

1. Interview design and data
   How do we elicit and record responses?

2. Validation
   Do responses reflect the public persona?

3. Cross-sectional return predictability
   Do responses predict future stock returns?

4. Beyond direction
   Disagreement, weak conviction, and macro evidence.

How do twin responses become stock-level signals?

Stock A · Tuesday

Ten fictional digital-twin responses



A	B	C	D	E	F	G	H	I	J
75	65	25	50	50	50	50	50	50	50

2 buys (≥ 65) | 1 sell (≤ 35) | 7 neutral (50)

Net Buy Share

Buy share minus sell share

$$\frac{\#\,\text{buys} - \#\,\text{sells}}{\text{all responses}} = \frac{2-1}{10} = 10\%$$

How do twin responses become stock-level signals?

Stock A · Tuesday

Ten fictional digital-twin responses

75	65	25	50	50	50	50	50	50	50

Net Buy Share	Buy share minus sell share $(2 - 1)/10 = 10\%$
Tilt	Mean of all scores, minus 50 $(25 + 15 - 25)/10 = 1.5$ points
Conviction-weighted meaningful Tilt	Mean meaningful tilt $\times$ meaningful share Meaningful: score ≥ 65 or ≤ 35 $[(25 + 15 - 25)/3] \times (3/10) = 1.5$

Account–stock–day responses become stock–day signals.

Fama–MacBeth regressions

Within each date, we compare stocks using the aggregated interview signals.

$$R^{e}_{j,t\to t+h} = a_t + b_t\, S_{j,t} + \varepsilon_{j,t}$$

Future h-day excess return relative to the S&P 500

Within-date association between signal and return

Signal: Net Buy Share, Tilt, CWM-Tilt.

Run the stock cross section on each event date using future returns.

Date 1	Date 2	⋯	Date T		Average
b_1	b_2	⋯	b_T	⟶	$\bar{b} = (b_1 + \cdots + b_T)/T$
					Fama–MacBeth estimate

Look at the average slope across the 50 stock interview dates.

Fama–MacBeth regressions

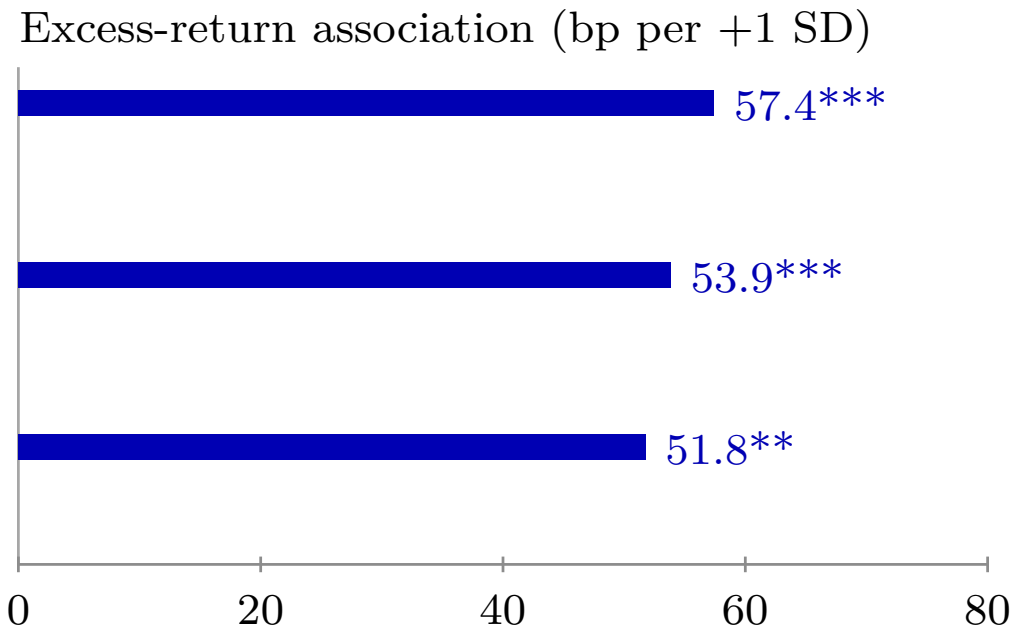
$$R^{e}_{j,t\to t+10} = a_t + b_t\, S_{j,t} + \varepsilon_{j,t}$$

Baseline (10 days)

Net Buy Share
(1 SD = 14.23 pp)

Tilt
(1 SD = 3.38 points)

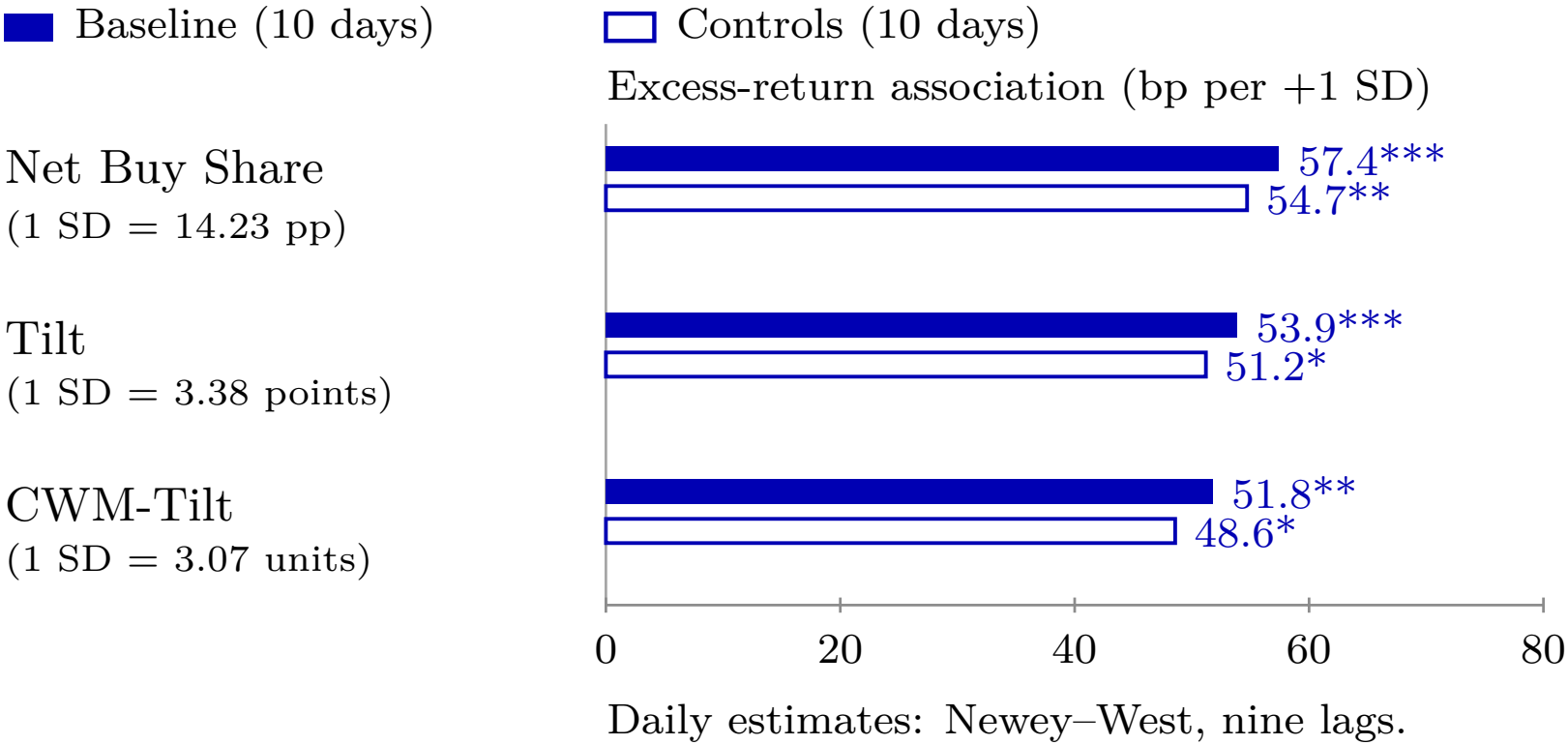
CWM-Tilt
(1 SD = 3.07 units)



Daily estimates: Newey–West, nine lags.

Fama–MacBeth regressions

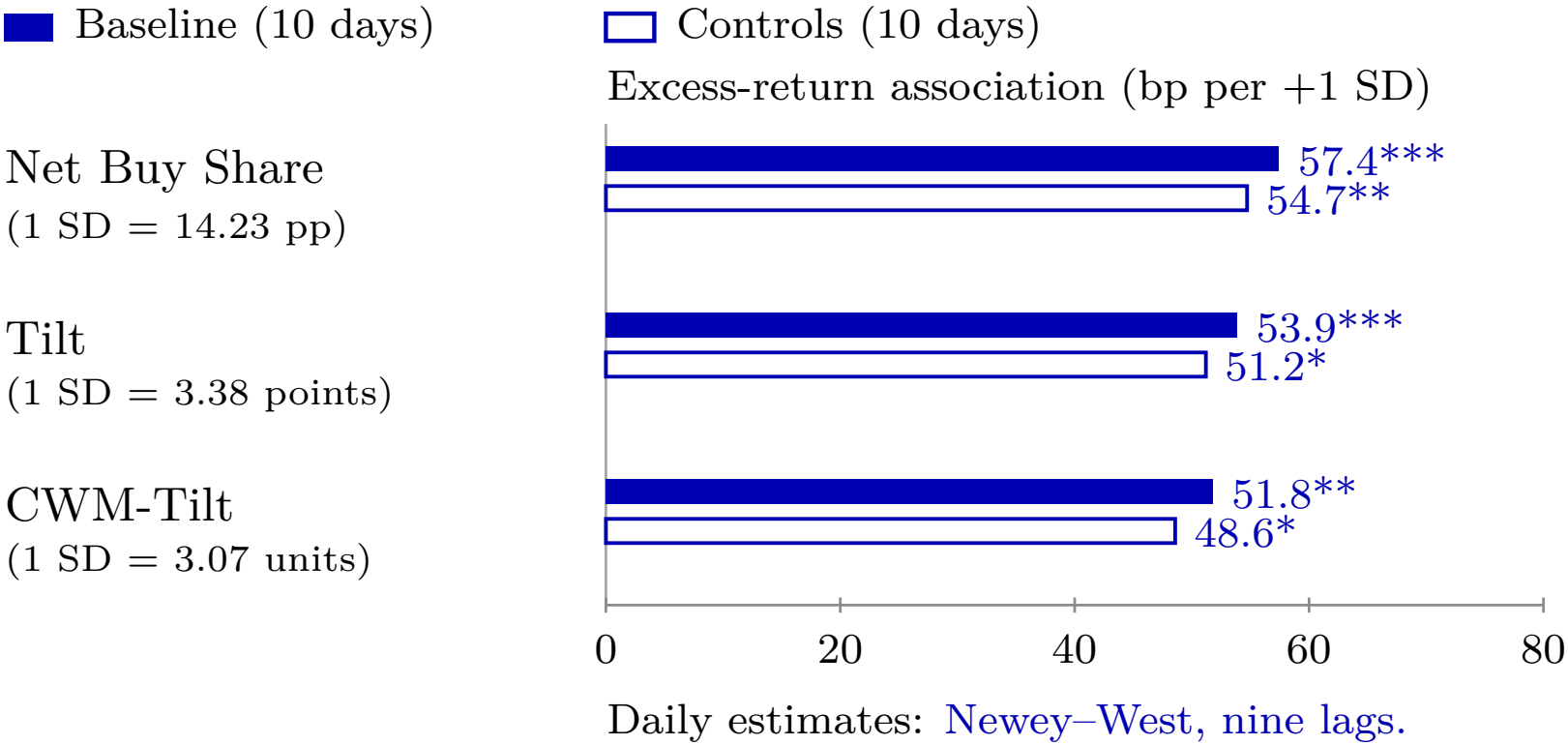
$$R^{e}_{j,t\to t+10} = a_t + b_t\, S_{j,t} + c_t' X^{\text{past}}_{j,t} + \varepsilon_{j,t}$$



Controls: prior ten-day excess return and realized excess-return volatility.

Fama–MacBeth regressions

$$R^{e}_{j,t\to t+10} = a_t + b_t\, S_{j,t} + c_t' X^{\text{past}}_{j,t} + \varepsilon_{j,t}$$

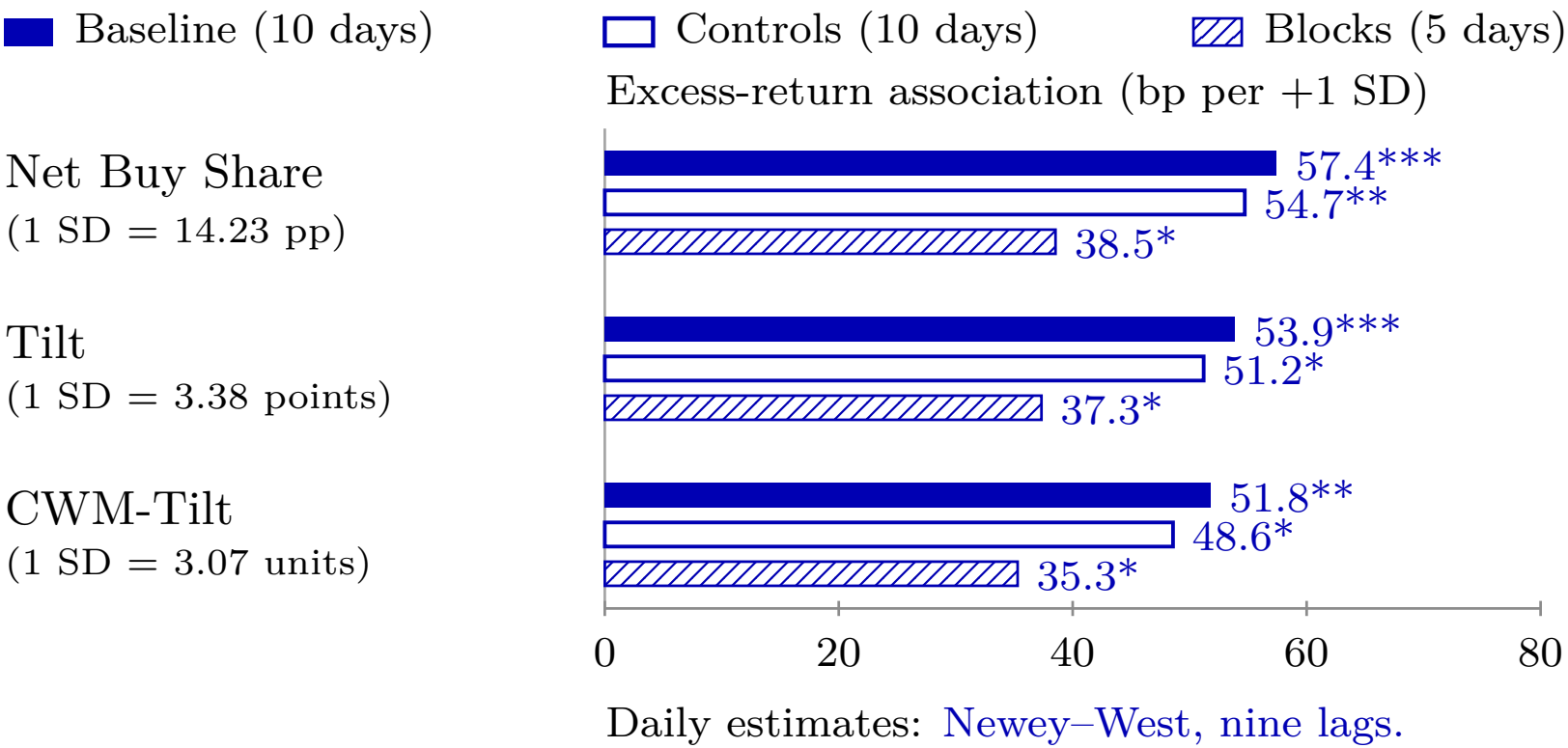


Controls: prior ten-day excess return and realized excess-return volatility.

Overlap: Newey–West adjusts for dependence in daily estimates.

Fama–MacBeth regressions

Non-overlapping five-day blocks



Daily estimates: Newey–West, nine lags.

Overlap: Wed–Tue signal averages predict the next block's return.

Digital-twin signals contain information about future returns.

Staggered ten-day return windows

Alignment 1

March 2026

Mon	Tue	Wed	Thu	Fri	Sat	Sun
(2)	3	4	5	6	7	8
9	10	11	12	13	14	15
(16)	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

At each circle, average recent signals up to this date.

The shaded future-return windows do not overlap.

Measure again ten trading days later.

◯ Measurement date Successive ten-day return windows

Staggered ten-day return windows

Alignment 1

March 2026

Mon	Tue	Wed	Thu	Fri	Sat	Sun
(2)	3	4	5	6	7	8
9	10	11	12	13	14	15
(16)	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

Alignment 2

March 2026

Mon	Tue	Wed	Thu	Fri	Sat	Sun
2	(3)	4	5	6	7	8
9	10	11	12	13	14	15
16	(17)	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

◯ Measurement date Successive ten-day return windows

Ten alignments. Return windows do not overlap within each.

Staggered ten-day return windows

Run one regression for each alignment, pooling its stock–date observations.

$$R^{e}_{j,t\to t+10} = a_{t,v} + b_v\,\overline{S}^{10}_{j,t} + c_v' X^{\text{past}}_{j,t} + \varepsilon_{j,t}$$

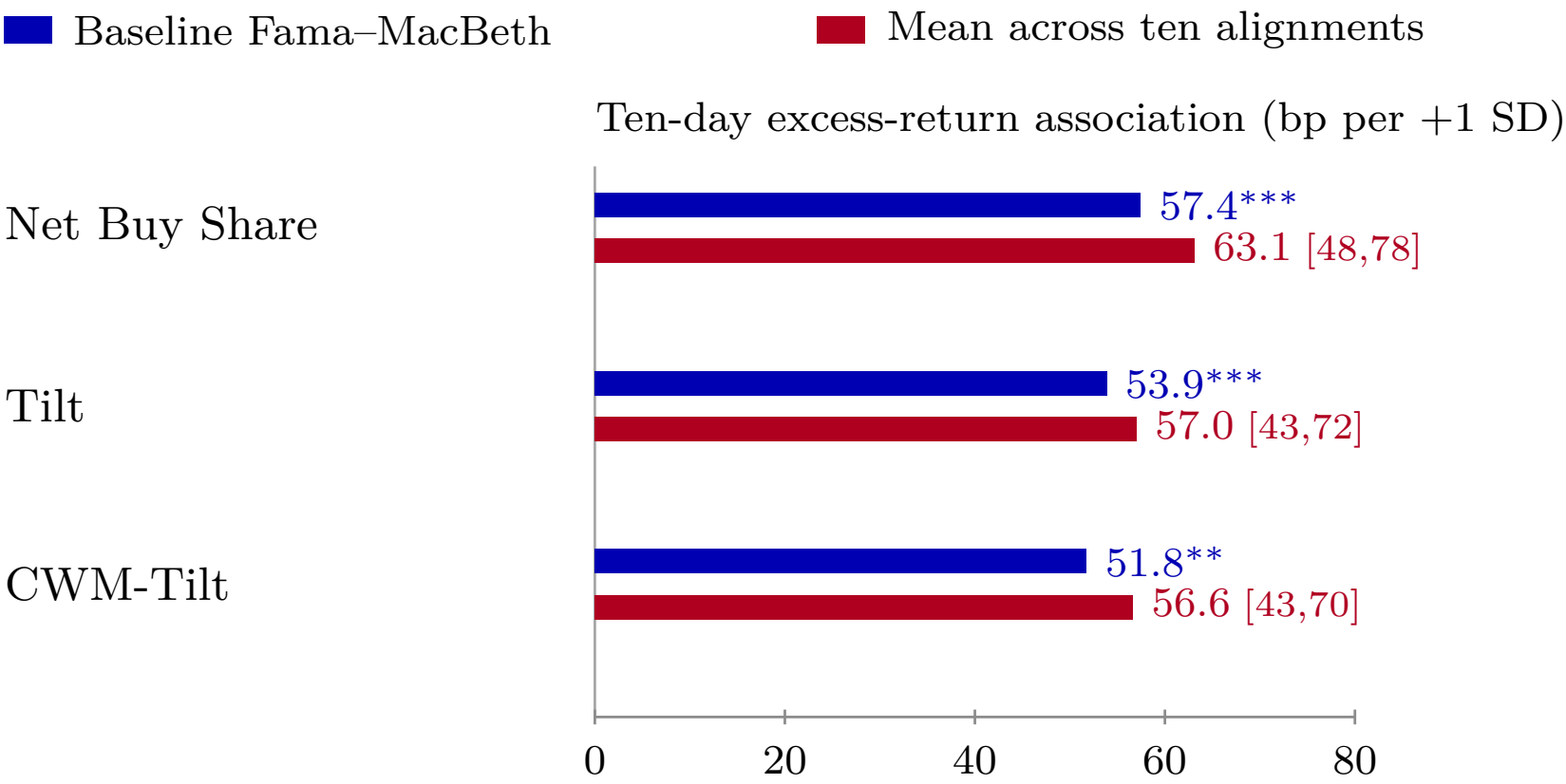
Signal	Ten-day mean ending on the measurement date
Controls	Prior excess return and realized excess-return volatility

Alignment 1 $\to b_1$ Alignment 2 $\to b_2$ $\cdots$ Alignment 10 $\to b_{10}$

How many estimates are positive? How many are significant?

Mean $N \approx$ 2,660: about 429 stocks $\times$ 6–7 windows.

Staggered ten-day return windows



10/10 alignment slopes are positive and significant at 10%.

Inference for the staggered-vintage estimates

The alignments share dates, and signals persist over time.

Could reassigned histories generate a mean this large?

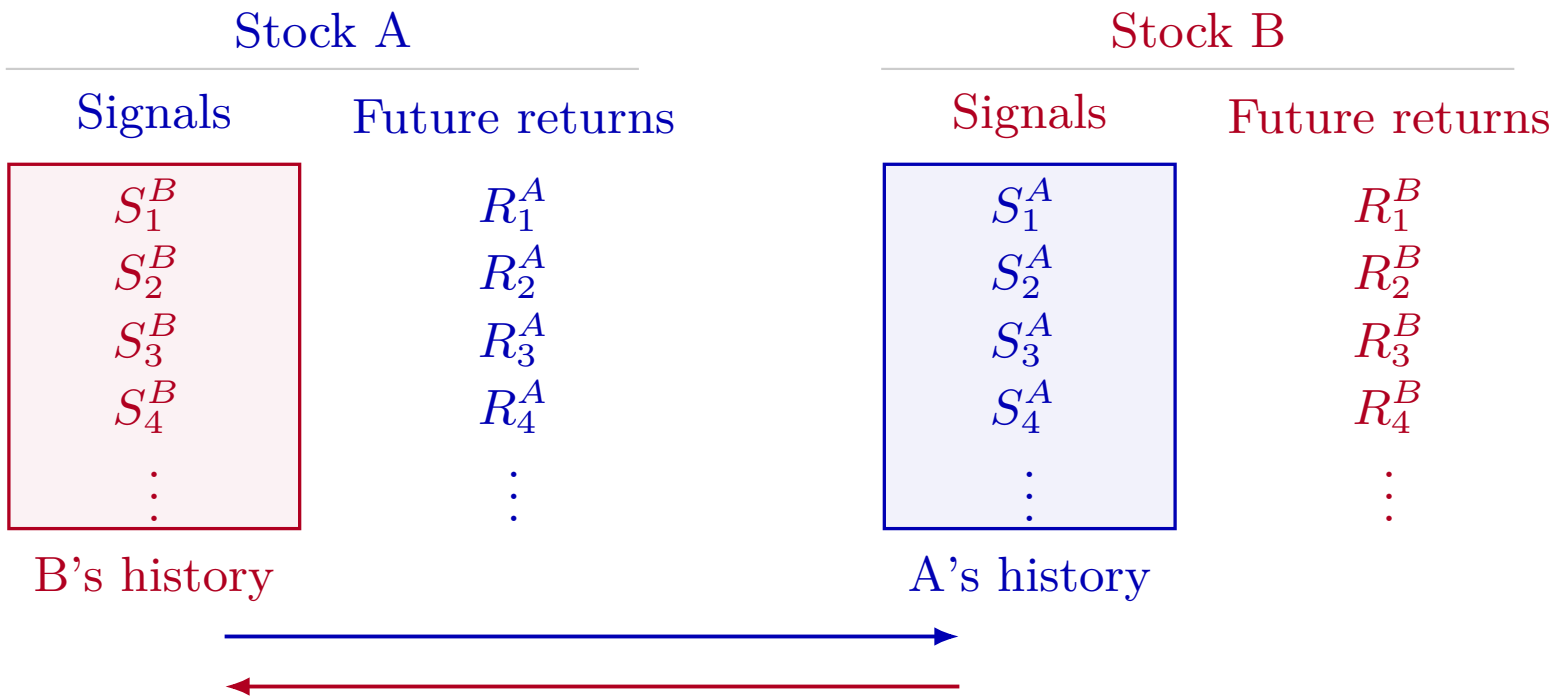
Stock A		Stock B	
Signals	Future returns	Signals	Future returns
S_1^A	R_1^A	S_1^B	R_1^B
S_2^A	R_2^A	S_2^B	R_2^B
S_3^A	R_3^A	S_3^B	R_3^B
S_4^A	R_4^A	S_4^B	R_4^B
$\vdots$	$\vdots$	$\vdots$	$\vdots$
A's history		B's history	

Keep stock returns and controls fixed.

Inference for the staggered-vintage estimates

The alignments share dates, and signals persist over time.

Could reassigned histories generate a mean this large?

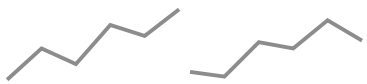


Re-estimate all ten alignments; save their mean slope $\bar{b}^*$.

Repeat 10,000 times; compare the observed mean with the placebos.

Inference for the staggered-vintage estimates

Four null hypotheses for the same observed return association.

Unrestricted	Any stock A ⟷ B	The observed relation is not unusual under random stock–signal matching.
AR simulation		The observed relation is not unusual given the signals' distributions and persistence.
Size quartiles	Initial size quartile A ⟷ B	The observed relation is not unusual once broad size-related patterns are preserved.
FF12 industry	Same industry A ⟷ B	The observed relation is not unusual once shared industry patterns are preserved.

Is the observed mean unusual under each benchmark?

How unusual is the observed return signal?

Net Buy Share (+1 SD): ten-day returns

Observed mean: 63.1 bp



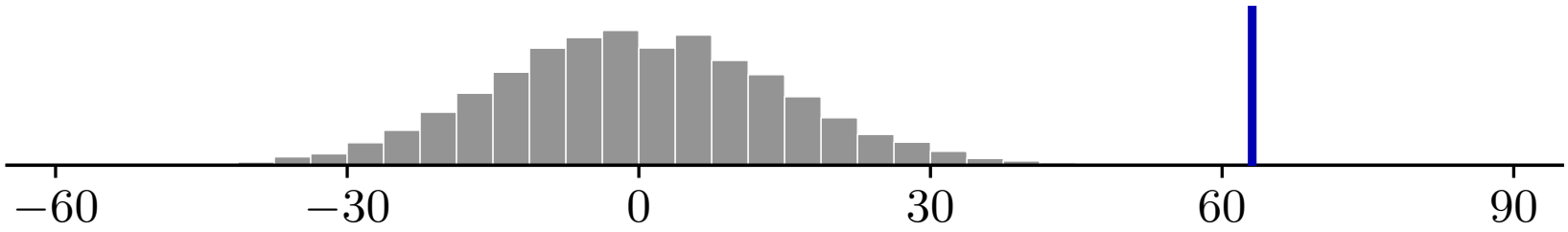
How unusual is the observed return signal?

Net Buy Share (+1 SD): ten-day returns

Observed mean: 63.1 bp

Unrestricted reassignment

$p = 0.0002$



Unusually informative relative to random reassignment.

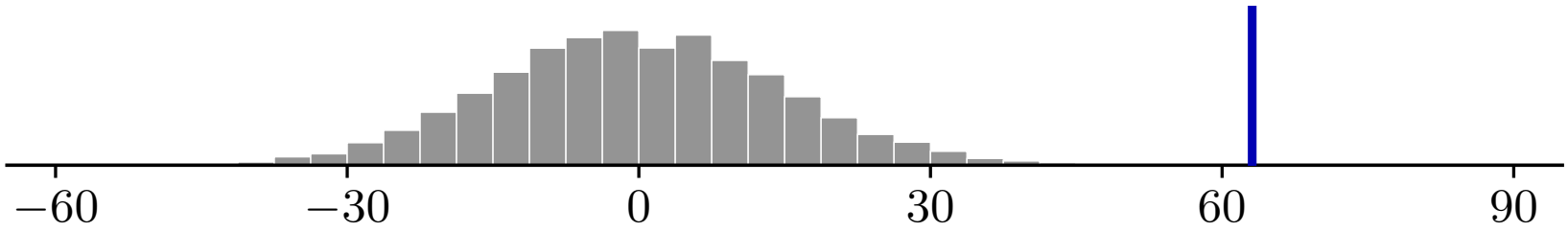
How unusual is the observed return signal?

Net Buy Share (+1 SD): ten-day returns

Observed mean: 63.1 bp

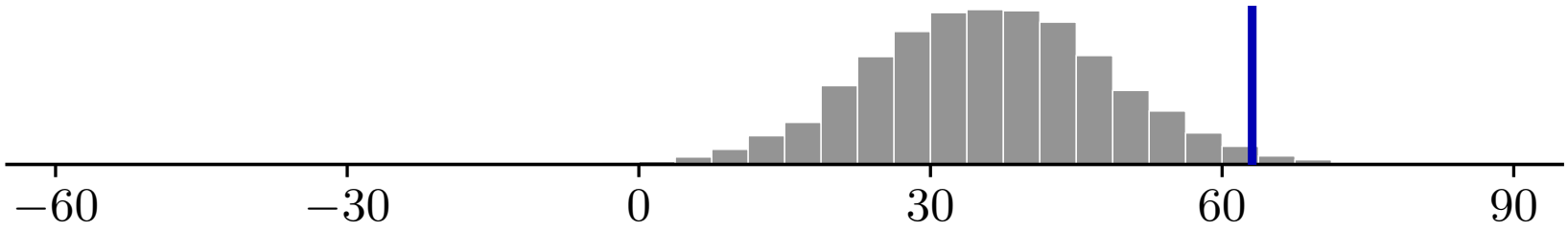
Unrestricted reassignment

$p = 0.0002$



Within-industry reassignment

$p = 0.0268$

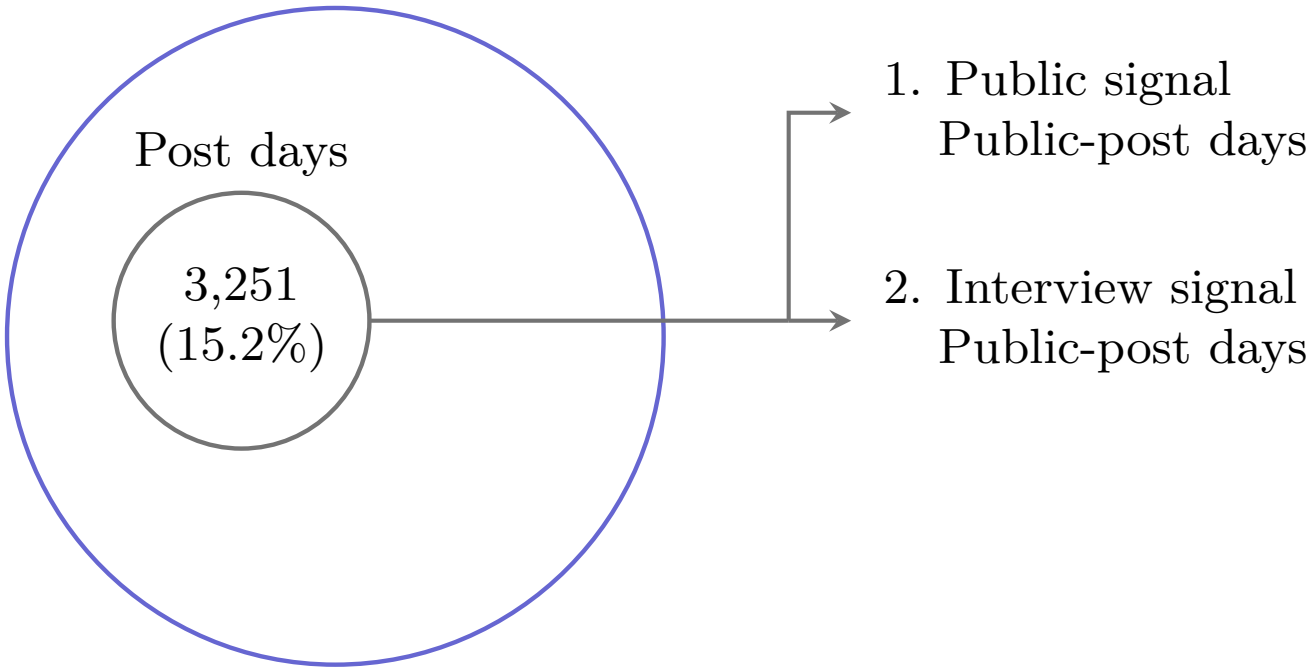


Industry matters, but does not fully explain the signal.

What about the silent region specifically?

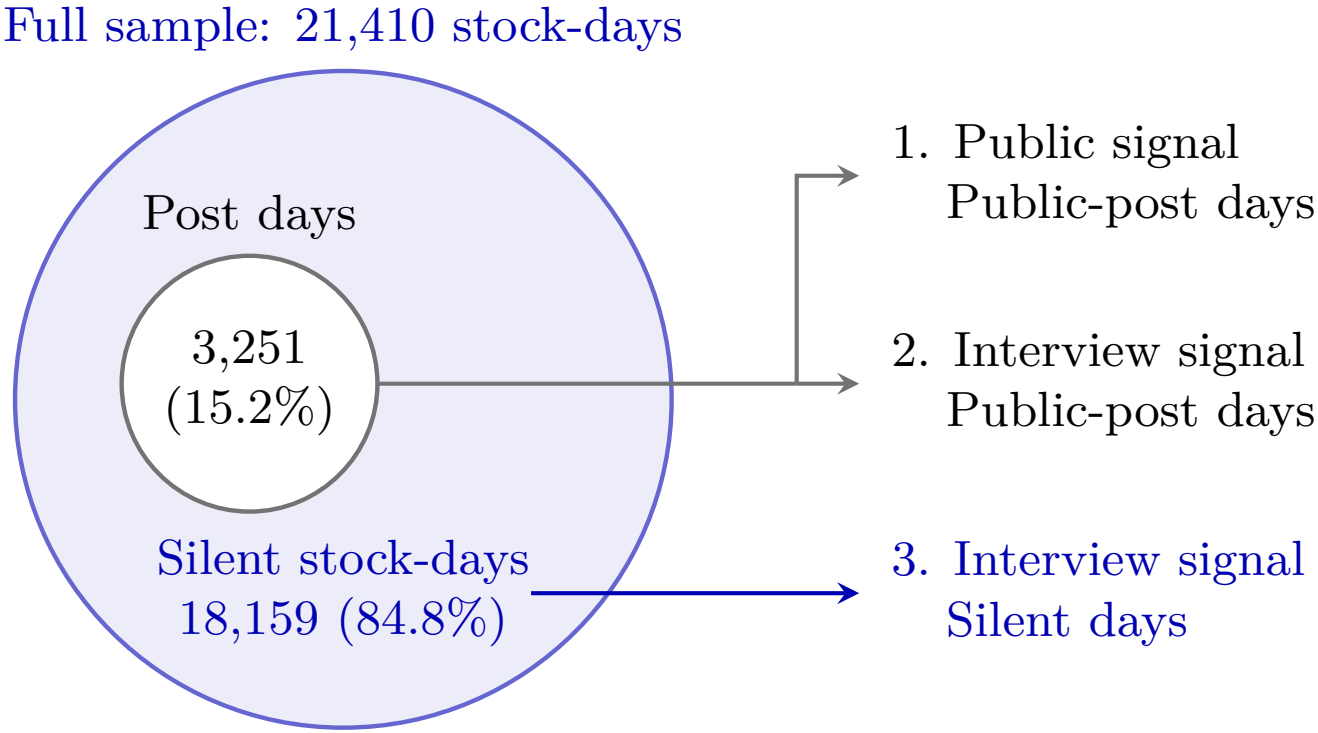
Partition the original stock-day panel by public recommendations.

Full sample: 21,410 stock-days



What about the silent region specifically?

Partition the original stock-day panel by public recommendations.

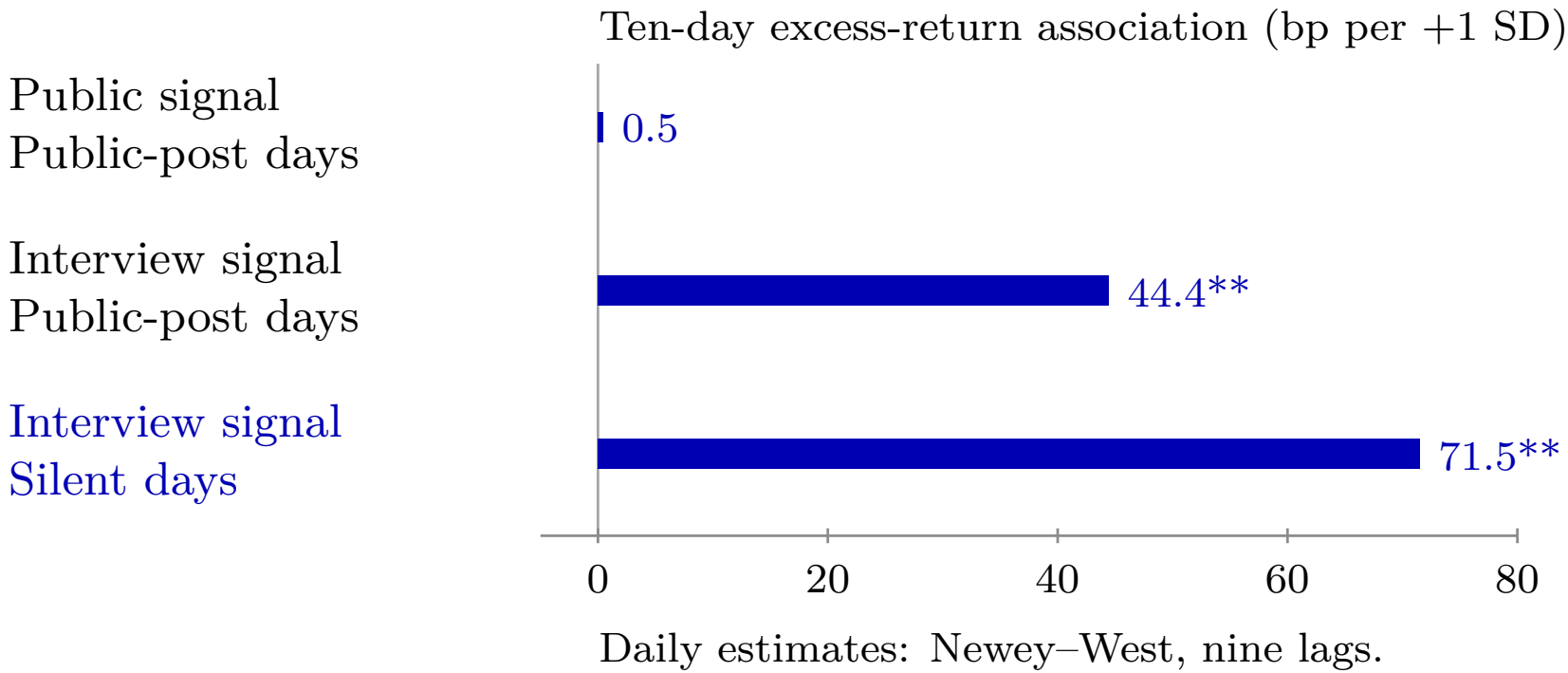


Do interviews predict returns when there is silence?

The additional coverage contains return information

$$R^{e}_{j,t\to t+10} = a_t + b_t\, S_{j,t} + c_t' X^{\text{past}}_{j,t} + \varepsilon_{j,t}$$

Net Buy Share: +1 SD

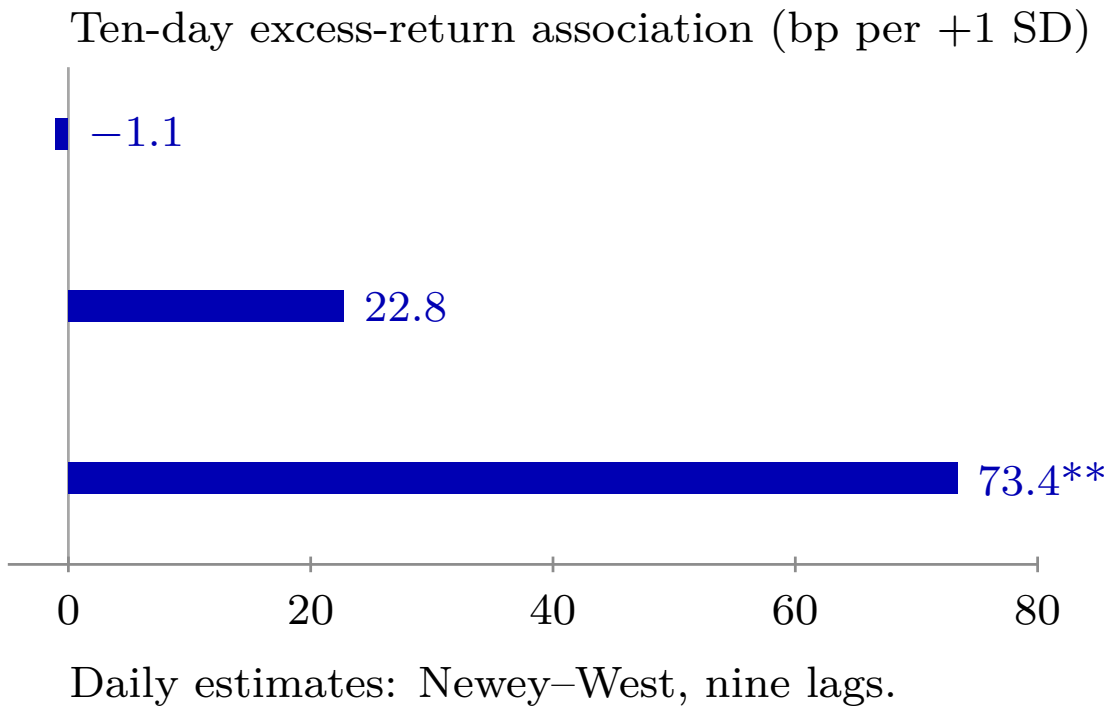


The silent region contains return information.

The additional coverage contains return information

$$R^{e}_{j,t\to t+10} = a_t + b_t\, S_{j,t} + c_t' X^{\text{past}}_{j,t} + \varepsilon_{j,t}$$

CWM-Tilt: +1 SD



The silent region contains return information.

Roadmap

1. Interview design and data
   How do we elicit and record responses?

2. Validation
   Do responses reflect the public persona?

3. Cross-sectional return predictability
   Do responses predict future stock returns?

4. Beyond direction
   Disagreement, weak conviction, and macro evidence.

Disagreement differs from uncertainty

Grounded disagreement



	A	B	C	D
Score	75	75	25	25
Speculation	20	20	20	20

Informative polarization

Weak conviction



	A	B	C	D
Score	50	50	75	25
Speculation	20	20	90	90

Uncertainty score

Same average recommendation (50). Different information.

How do disagreement and uncertainty predict returns?

Fama–MacBeth regressions, as before.

$$Y_{j,t\to t+10} = a_t + b_t \,\mathrm{State}_{j,t} + c_t' X_{j,t}^{\mathrm{past}} + \varepsilon_{j,t}$$

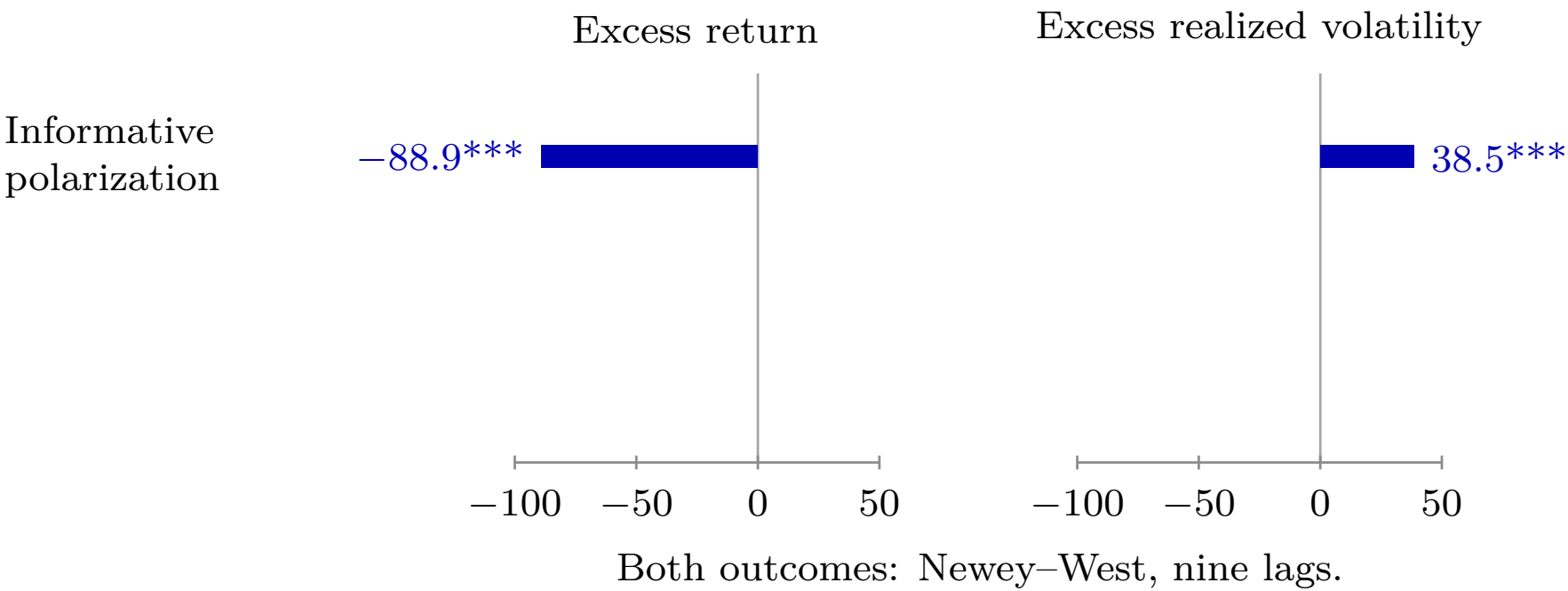
Ten-day excess return
or realized excess volatility

Informative polarization
or uncertainty score

How do disagreement and uncertainty predict returns?

$$Y_{j,t\to t+10} = a_t + b_t\,\mathrm{State}_{j,t} + c_t' X_{j,t}^{\mathrm{past}} + \varepsilon_{j,t}$$

Fama–MacBeth estimates (bp per +1 SD)



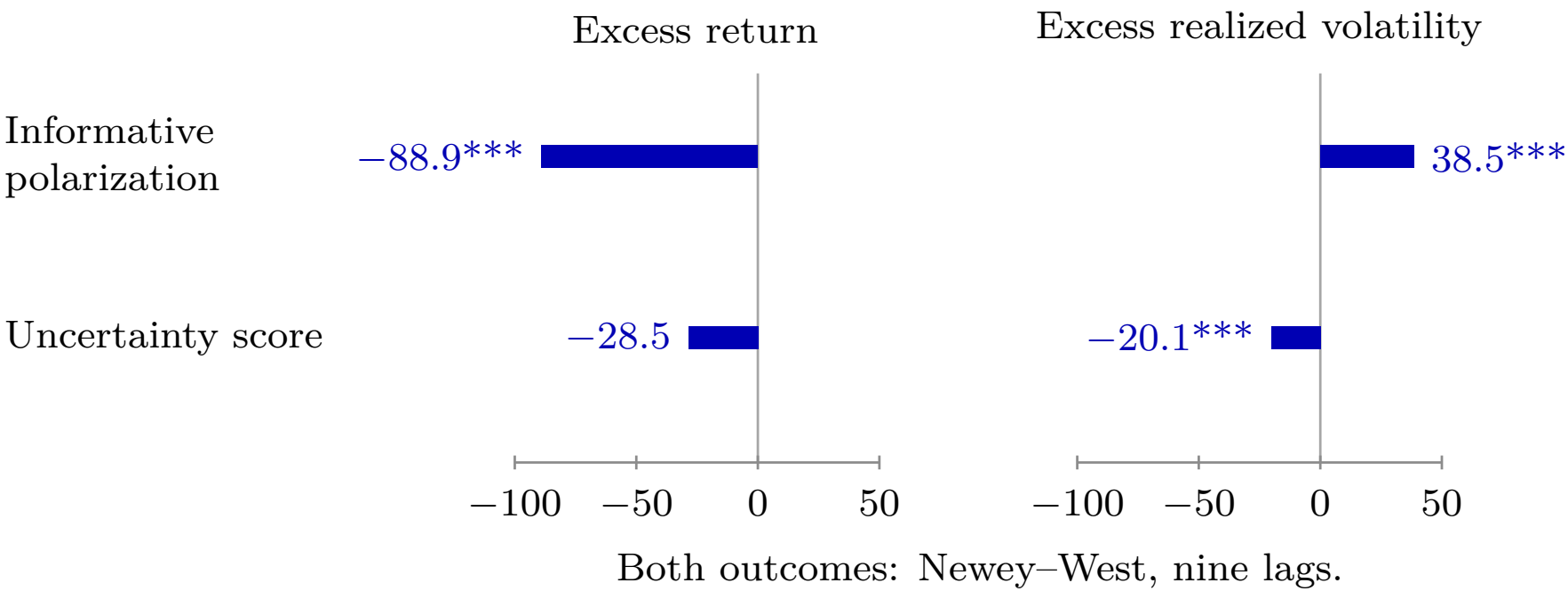
Both outcomes: Newey–West, nine lags.

Grounded disagreement predicts lower returns and higher volatility.

How do disagreement and uncertainty predict returns?

$$Y_{j,t\to t+10} = a_t + b_t\,\mathrm{State}_{j,t} + c_t' X_{j,t}^{\mathrm{past}} + \varepsilon_{j,t}$$

Fama–MacBeth estimates (bp per +1 SD)

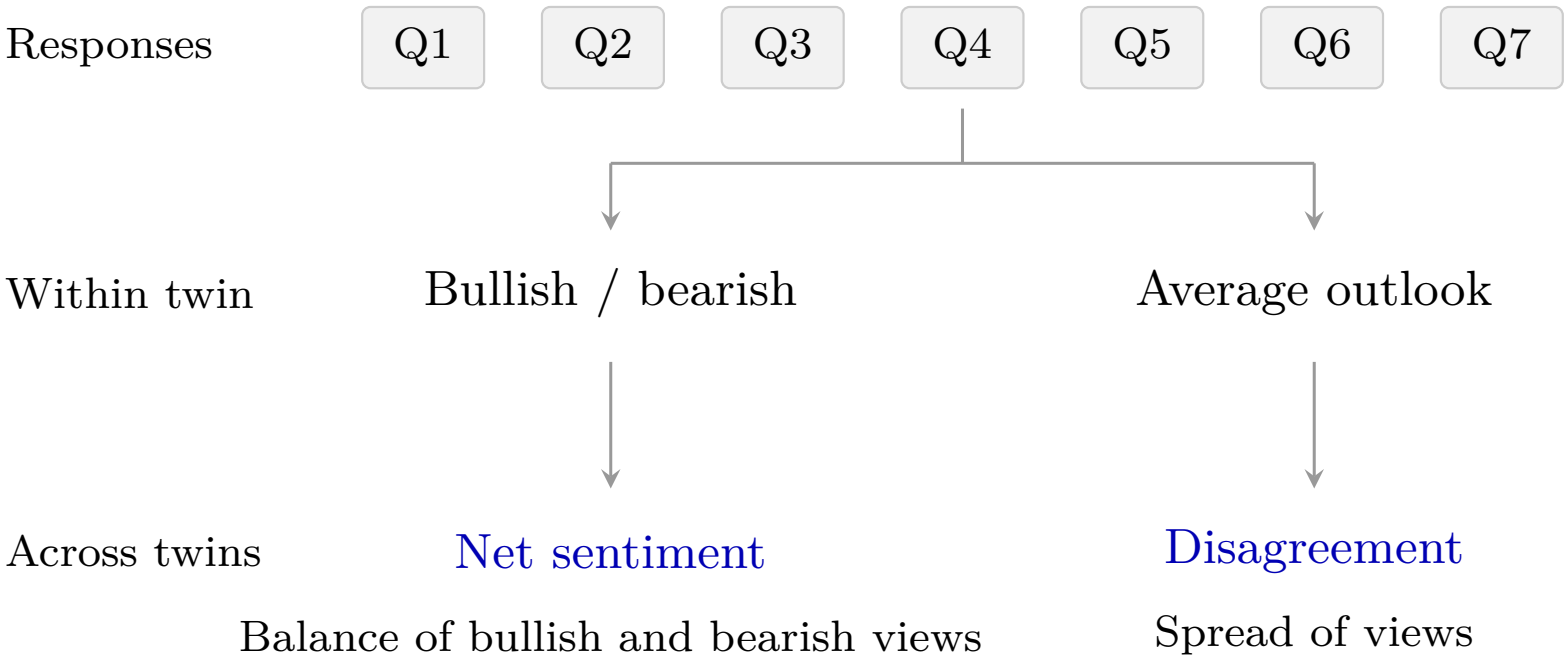


Both outcomes: Newey–West, nine lags.

Higher volatility with disagreement; lower with weak conviction.

What about all those macro questions?

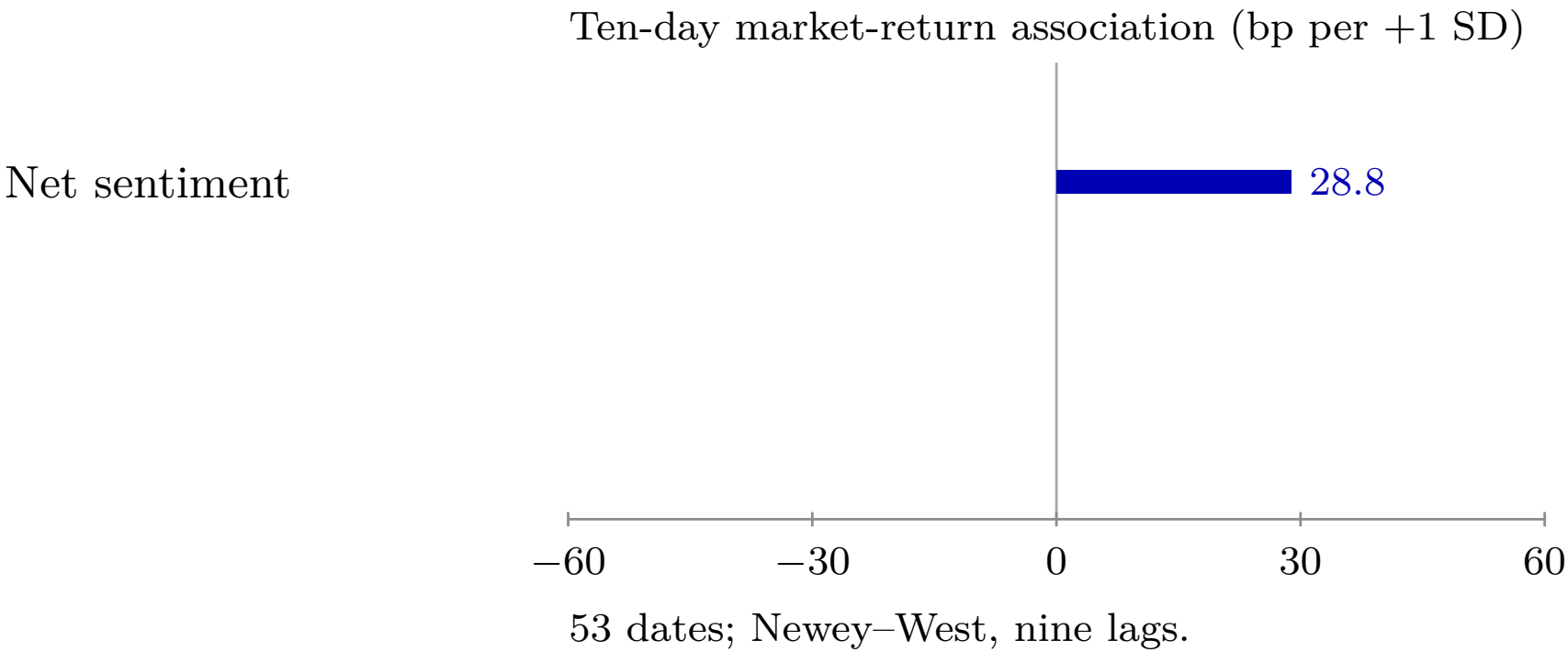
Summarize each twin's seven answers, then aggregate across twins.



Two market-level measures on each of 53 dates.

Do aggregate beliefs predict the market's next ten-day return?

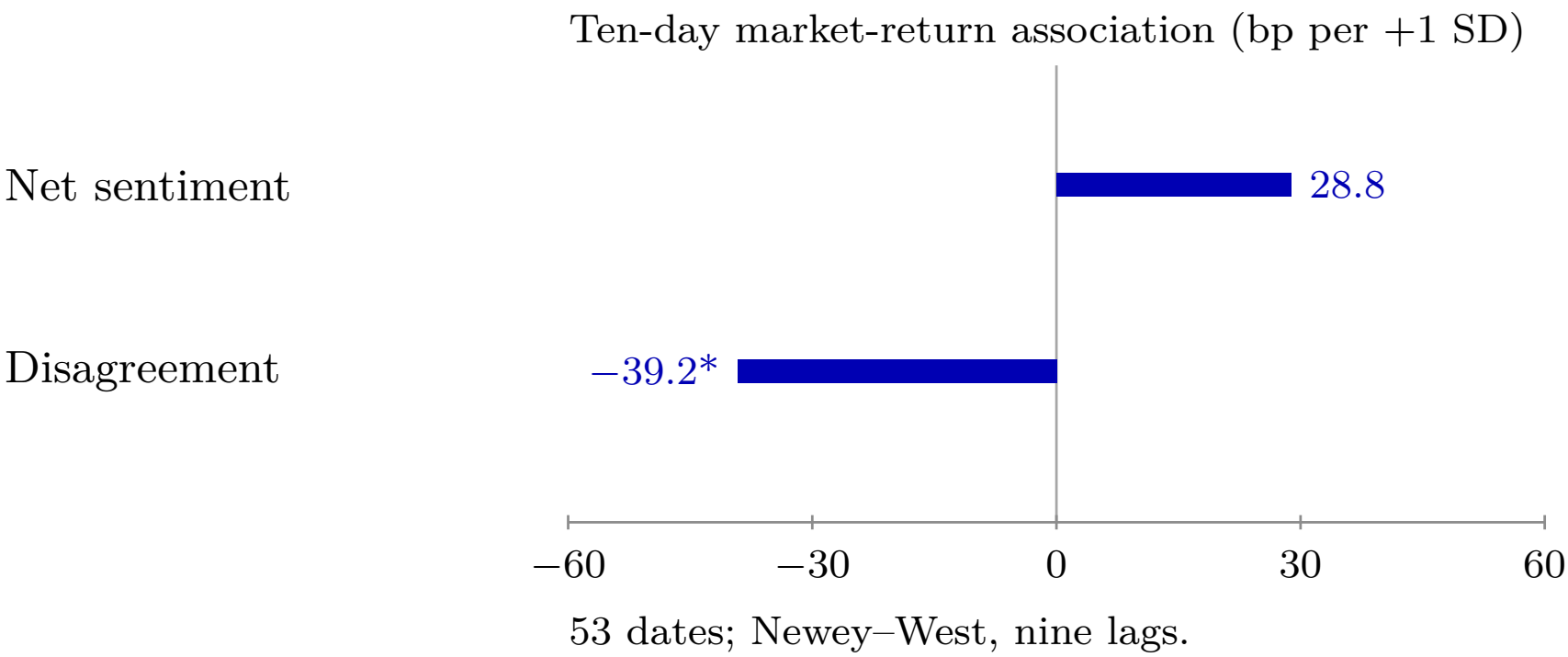
$$R^{\text{S\&P 500}}_{t \to t+10} = \alpha + \beta_S \, \text{Sentiment}_t + \gamma' X^{\text{past}}_t + \varepsilon_t$$



53 dates; Newey–West, nine lags.

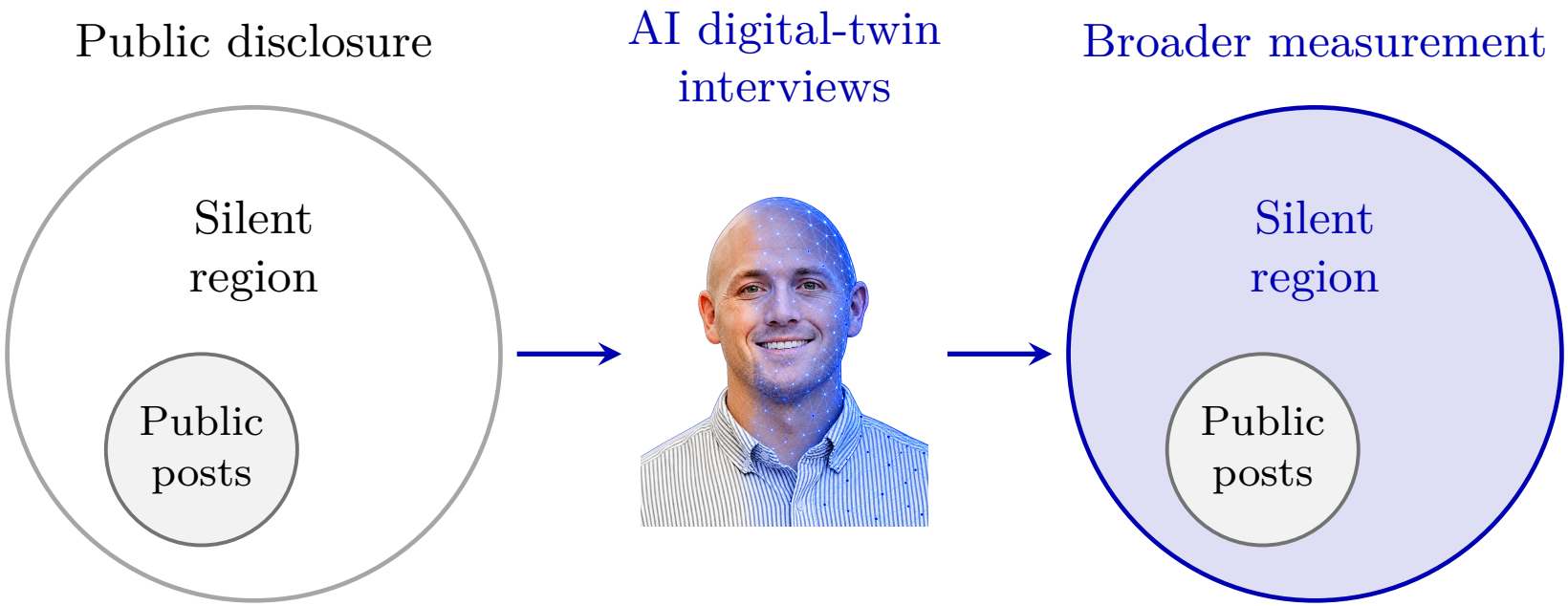
Do aggregate beliefs predict the market's next ten-day return?

$$R^{\text{S\&P 500}}_{t\to t+10} = \alpha + \beta_S\,\text{Sentiment}_t + \beta_D\,\text{Disagreement}_t + \gamma' X^{\text{past}}_t + \varepsilon_t$$



More macro disagreement is followed by lower market returns.

Coloring in the silent region



Digital twins give the silent region its voice.

The added coverage contains return information.

Thank you.