Class 12: Evidence from CAPM and APT IV

February 26, 2026

Paul Goldsmith-Pinkham
Yale University

https://paulgp.com

Yale SCHOOL OF MANAGEMENT

Today's Question

If factors explain returns, are they capturing real risks or exploitable opportunities?

How efficient are markets, really?

Positioning

Where We've Been	Where We Are	Where We're Going
Factor models: FF3, momentum, liquidity	Risks vs. opportunities, the factor zoo, and market efficiency (all three forms)	The behavioral critique and limits to arbitrage

Learning Objectives

By the end of today's class, you should be able to:

1. Articulate the debate over whether factor premia represent risk compensation or mispricing
2. Explain why the "factor zoo" is a problem and how machine learning methods address it
3. Define the three forms of the Efficient Market Hypothesis
4. Describe evidence on return predictability and seasonality under the weak form

5. Evaluate evidence for and against semi-strong and strong form efficiency

Factors: Risks or Opportunities?

Note the theme here:

- No shares are over/underpriced (almost)
- Risk-premia paid on assets represent exposure to risk factors
- Otherwise, “arbitrageurs” will quickly drive prices to equilibrium “correct” values
- They need deep pockets!

No free lunch $\longrightarrow$ high returns = high risk exposure

Different **kinds** of risk:

- Some institutions/investors prefer certain types of exposures

We discussed many potential factors: Fama-French 3, Momentum, Liquidity...

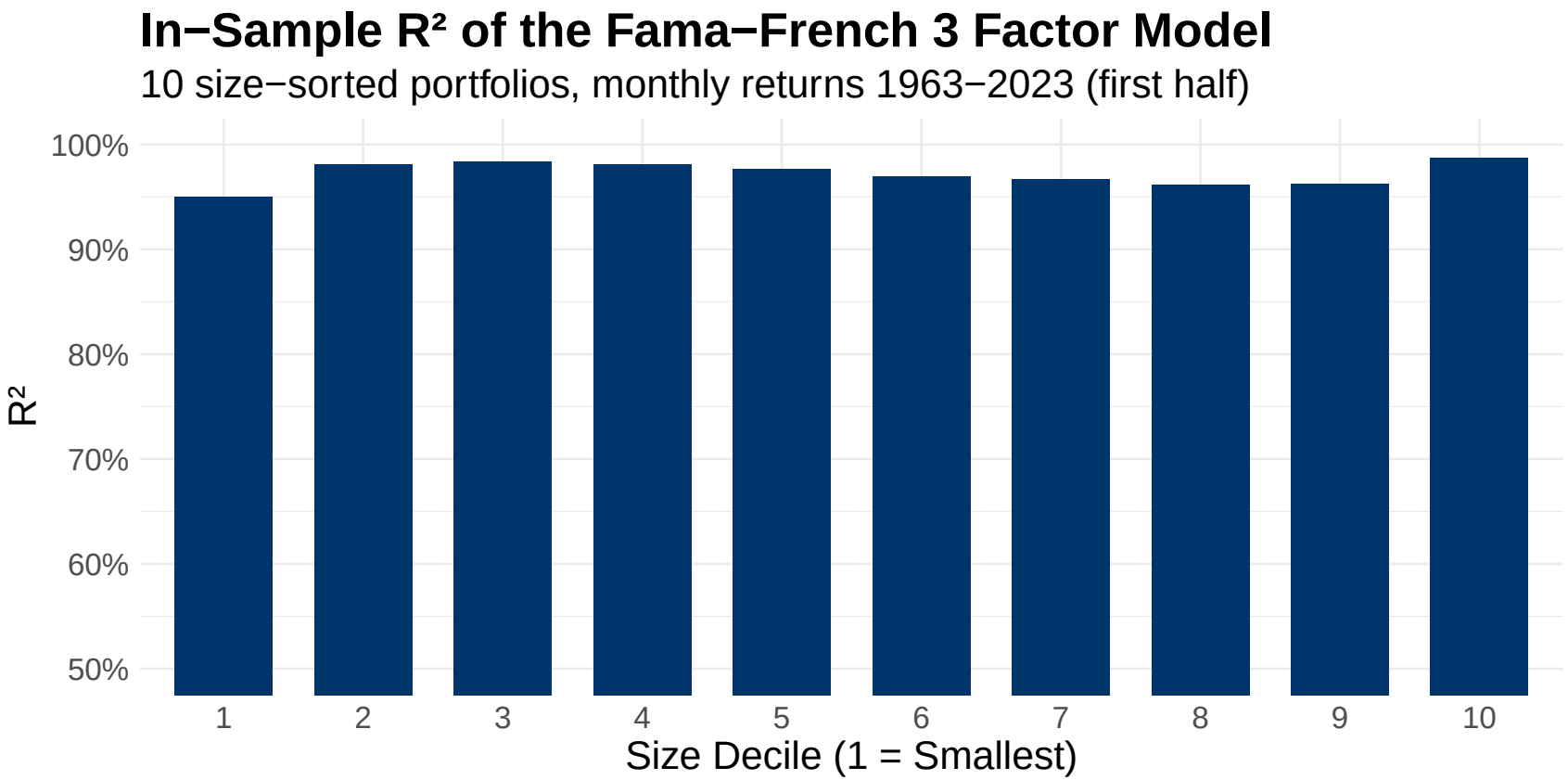
Why stop at 5? Can continue to capture as many risks as possible!

But what happens when we keep adding factors?

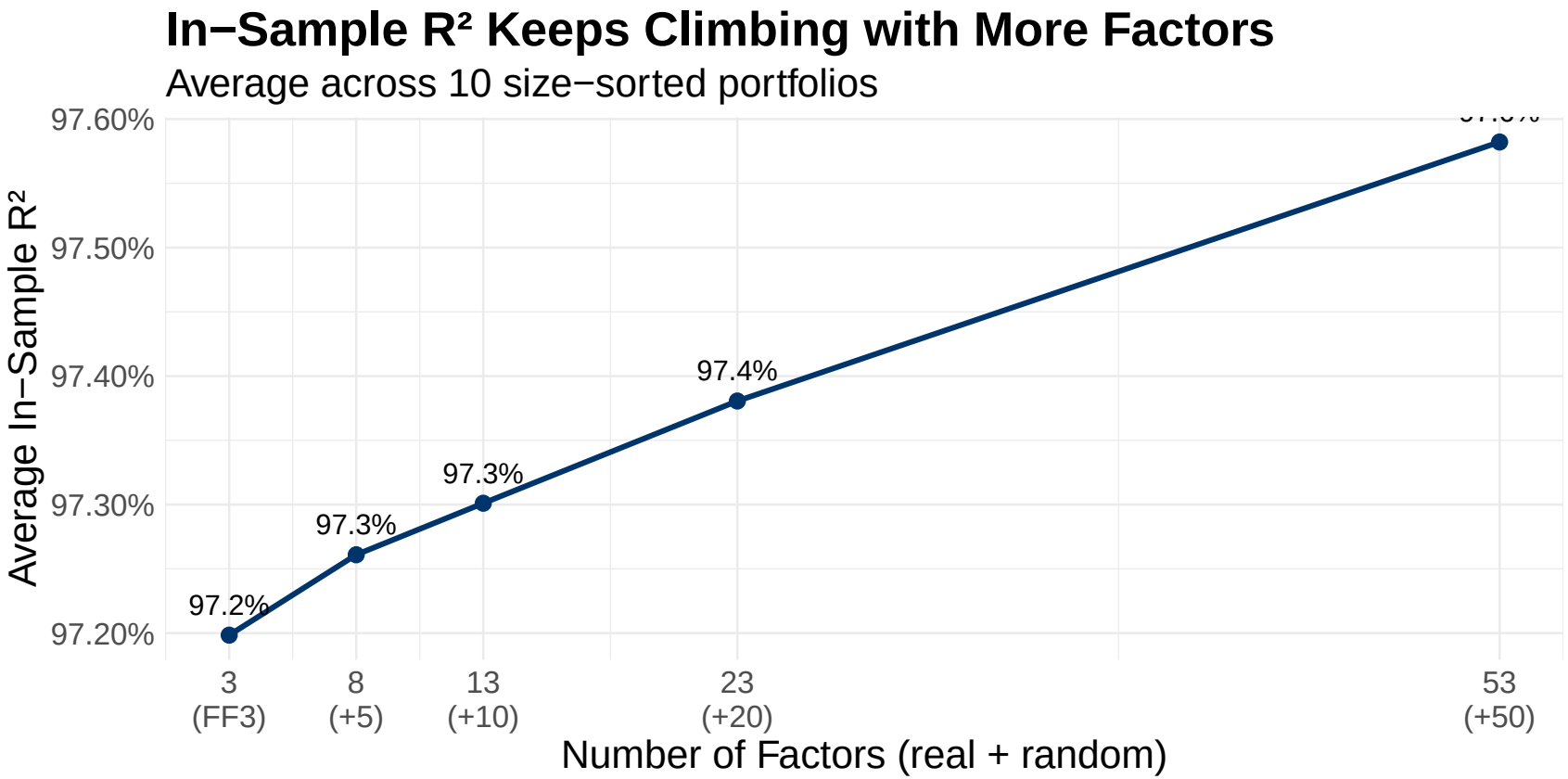
Let's run an experiment:

- Start with the FF3 model on 10 size-sorted portfolios (1963–2023)
- Progressively add 5, 10, 20, 50 *random noise* "factors"
- Estimate on the first half of the data, predict on the second half

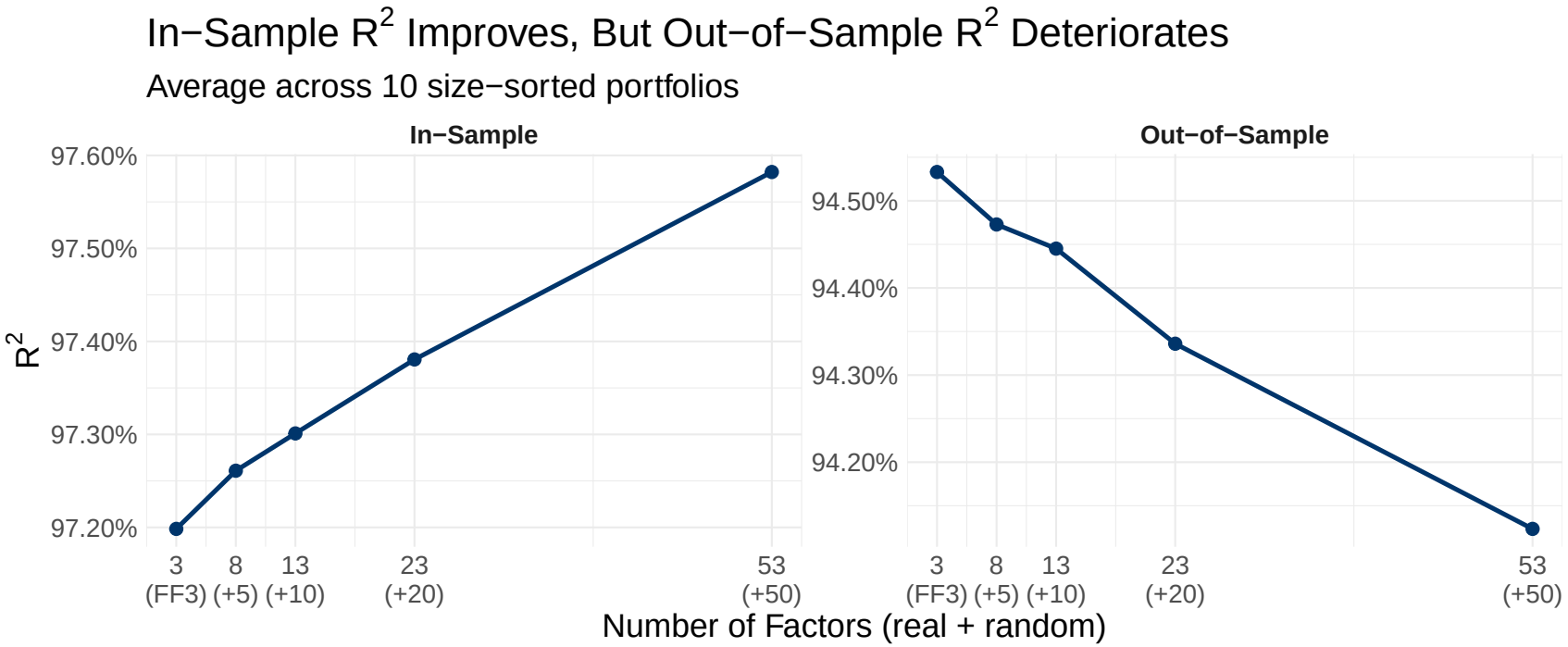
Starting Point: FF3 Works Well



More Factors = Better Fit... Right?



Garbage In, Garbage Out



The lesson: more factors $\neq$ better predictions.

Feng, Giglio and Xiu propose a solution: use “double-selection Lasso” to let the data identify which factors actually matter.

They test 15 new contributed factors from the literature — only 4 out of 15 are predictive beyond what was already known.

Key takeaway: Disciplined factor selection beats the kitchen sink.

Anomalies: Risks or Opportunities?

Note the theme:

- Shares are generally correctly priced
- Risk-premia paid on assets represent exposure to risk factors
- Otherwise, “arbitrageurs” will quickly drive prices to equilibrium

No free lunch $\longrightarrow$ high returns = high risk

These ideas are closely associated with the Efficient Market Hypothesis (EMH):

- Specifically, EMH posits that prices reflect all available information

The EMH comes in three flavors:

1. **Weak form:** all available information is limited to historical prices
2. **Semi-strong form:** all available data refers to all publicly available data
3. **Strong form:** all available information includes insider information

Examples (Weak Form)

Weak form: all available information is limited to historical prices

Empirical evidence shows that using historical data, we can predict:

1. Short-run reversals
2. Medium momentum
3. Long-term reversals

More recently: evidence on return seasonals (Heston and Sadka, 2008)

In the longer run, there is definitely return predictability...but that doesn't necessarily imply markets are inefficient.

Return Seasonality (Heston and Sadka 2008)

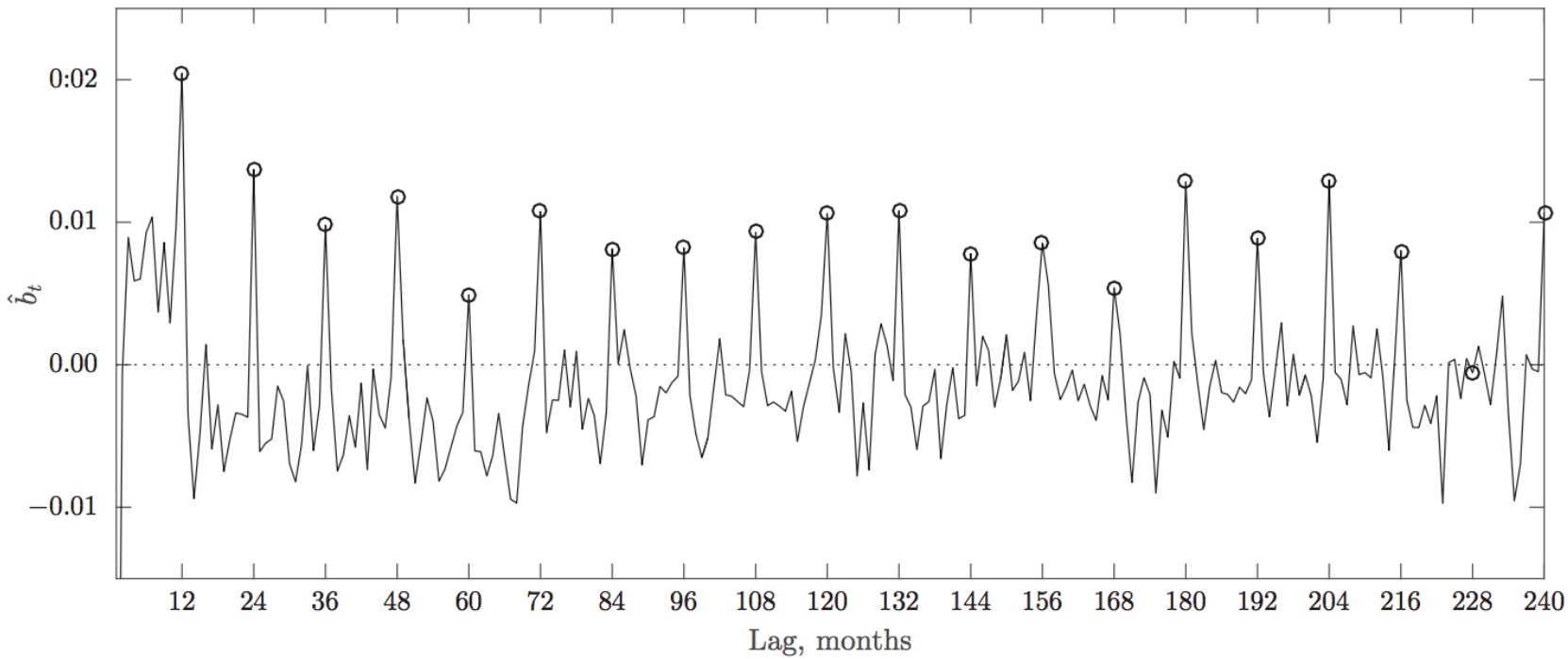
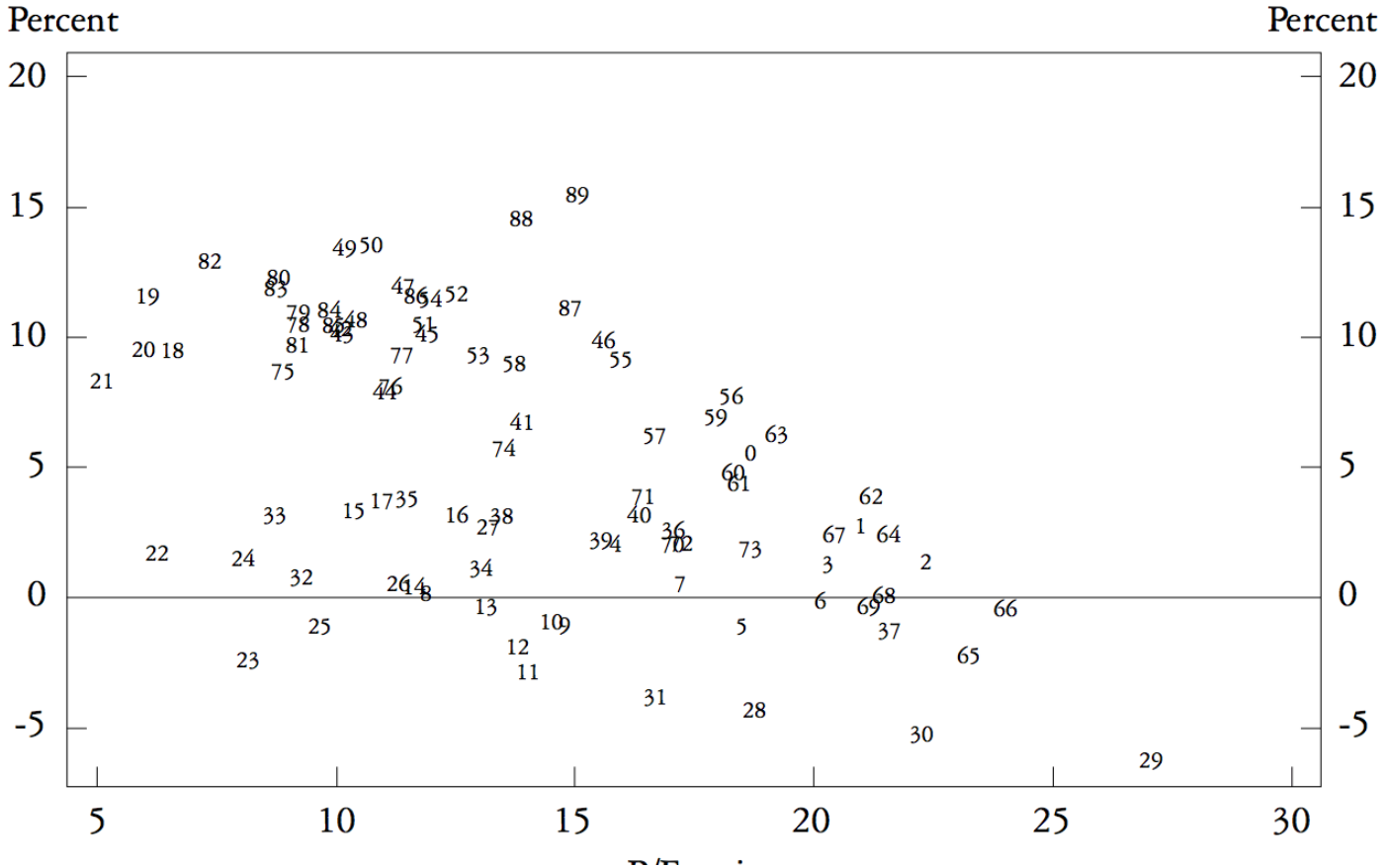


Figure 1. Seasonalities in individual stock returns. This figure plots slope coefficients from univariate Fama and MacBeth (1973) regressions of month t returns against month $t-k$ returns, $r_{i,t} = a_t + b_t\, r_{i,t-k} + e_{i,t}$, with k ranging from 1 to 240 months. The circles denote estimates at annual lags. The regressions use monthly data from January 1963 through December 2011 for NYSE, Amex, and NASDAQ stocks.

S&P 500 P/E Ratio

P/E RATIO AND STOCK PRICE GROWTH IN THE FOLLOWING 10 YEARS

Annual stock price growth



Examples (Semi-Strong)

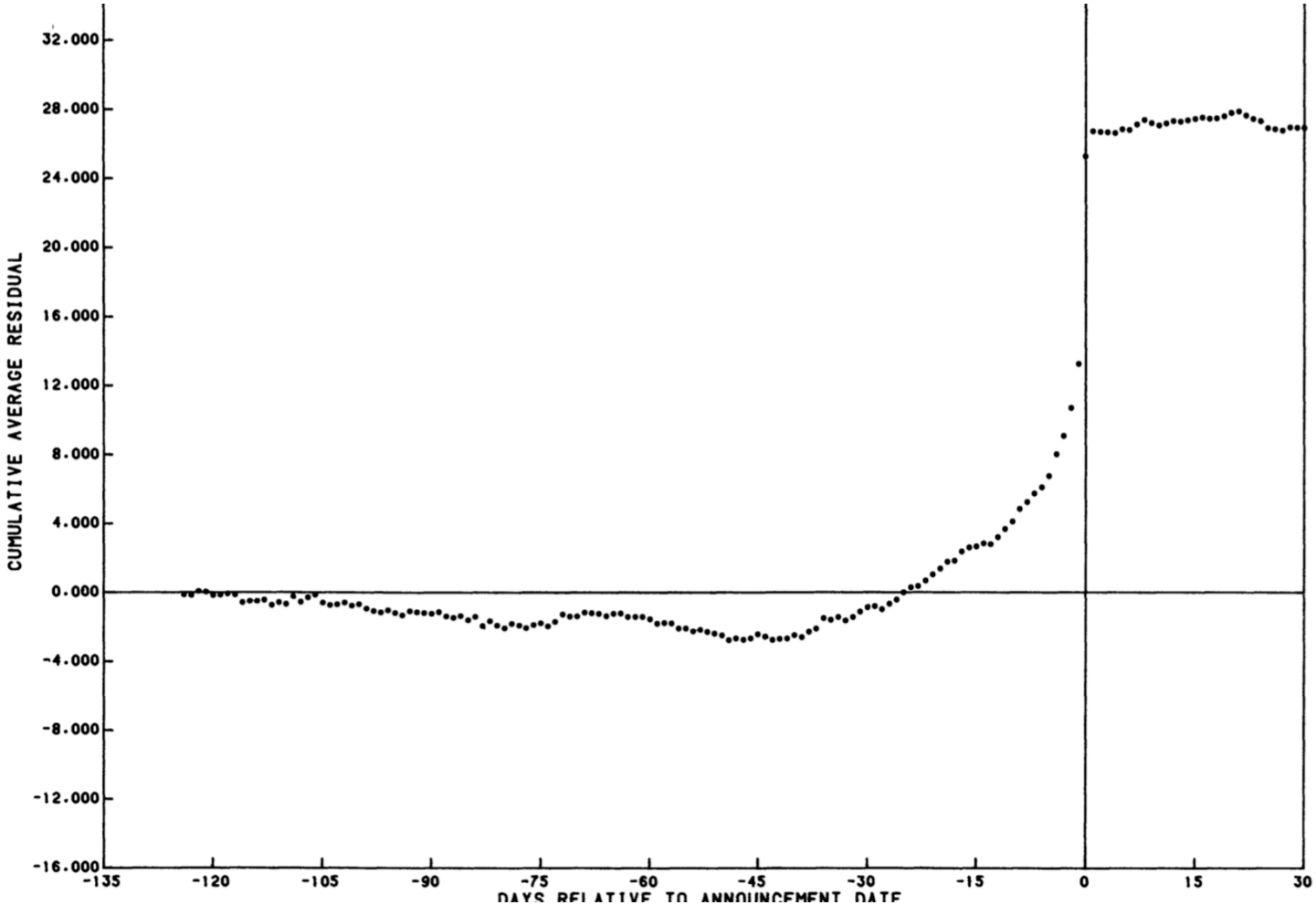
Da, Engelberg, and Gao (2010 and 2011) showed that it took time for markets to recognize value in new public information produced by Google trends.

Google search volume based on products predicted earnings better than both analyst forecasts and earnings announcement returns.

Search volume also proxied for retail investor interest:

- An increase predicts higher stock prices in the next 2 weeks and an eventual price reversal within the year
- It also contributes to the large first-day return and long-run underperformance of IPO stocks

Examples (Strong)



Next Class: March 3

Topics: Evidence from CAPM and APT V

- The behavioral critique
- Limits to arbitrage

Matt Levine Reading: TBD