BERD Academy & ifo Workshop "Novel Financial Data for Research"

# FINANCIAL MARKETS AND PRIVATE INVESTORS: DATA SOURCES, RESEARCH RESULTS, AND PERSPECTIVES AHEAD

Steffen Meyer
Professor
steffen.meyer@econ.au.dk

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

AARHUS
BSS

DANISH FINANCE
INSTITUTE

AACSB
ACCREDITED

ASSOCIATION
AMBA'S
ACCREDITED

EFMD
EQUIS
ACCREDITED

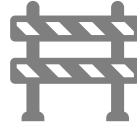**?**

**Research question**

What are the long-run returns to stocks in the US?



Lawrence Fischer
Associate
Professor of Finance



James H. Lorie
Professor

**🚧**

**Complication**

1) Data in physical tapes at the NYSE

2) Manual labor required

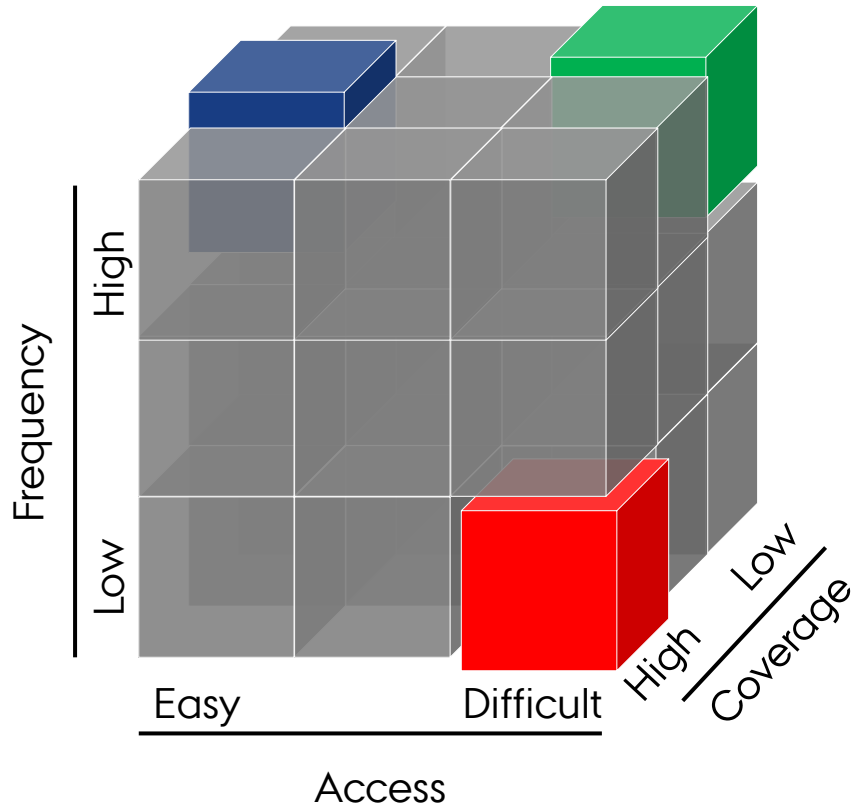3) Computing facilities are insufficient

**!**

**Solution**

Grant:



50.000 $

Impact

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

# When it comes to data (collection), we need to circle back to finance 101: There is no free lunch!



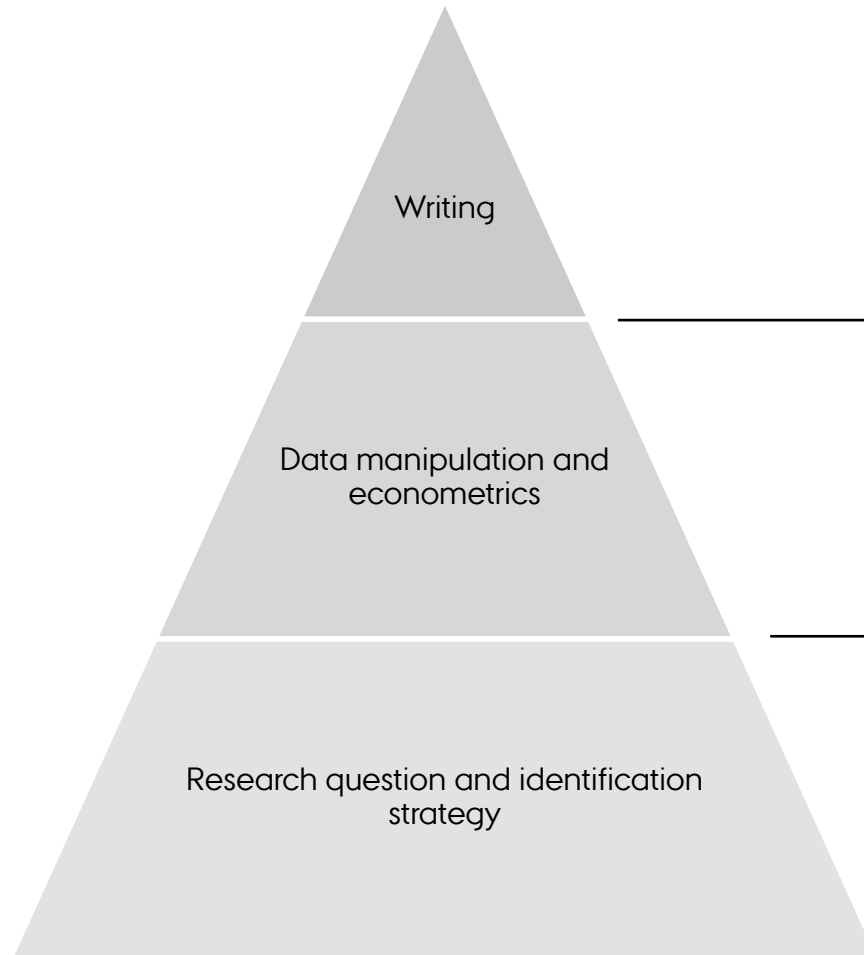| | Type | Biggest challenge |
|---|---|---|
| 🟥 (red) | Registry data (Example: Danish registry data) | • Access<br>• Define request<br>• Time / Costs |
| 🟦 (blue) | Commercial data (Example: WRDS, Bloomberg) | • Price<br>• Quotas<br>• Price |
| 🟩 (green) | Administrative data (Example: „Odean" data) | • GDPR<br>• Define request<br>• Quality control |

Prof. Dr. Steffen Meyer

# Data is not self-sufficient: It needs to be completed with a relevant question and a proper identification strategy to establish causality

**‚Kano'-pyramid of a research paper**



**Key questions:**
- Is the structure well described?
- Is the paper focusing on the one key message?
- Are there no spelling errors?

**Key questions:**
- Is the data acquisition well documented?
- Are code file replicable?
- Am I using the latest (most appropriate) method?

**Key questions:**
- Is the research novel?
- Is the research question relevant?
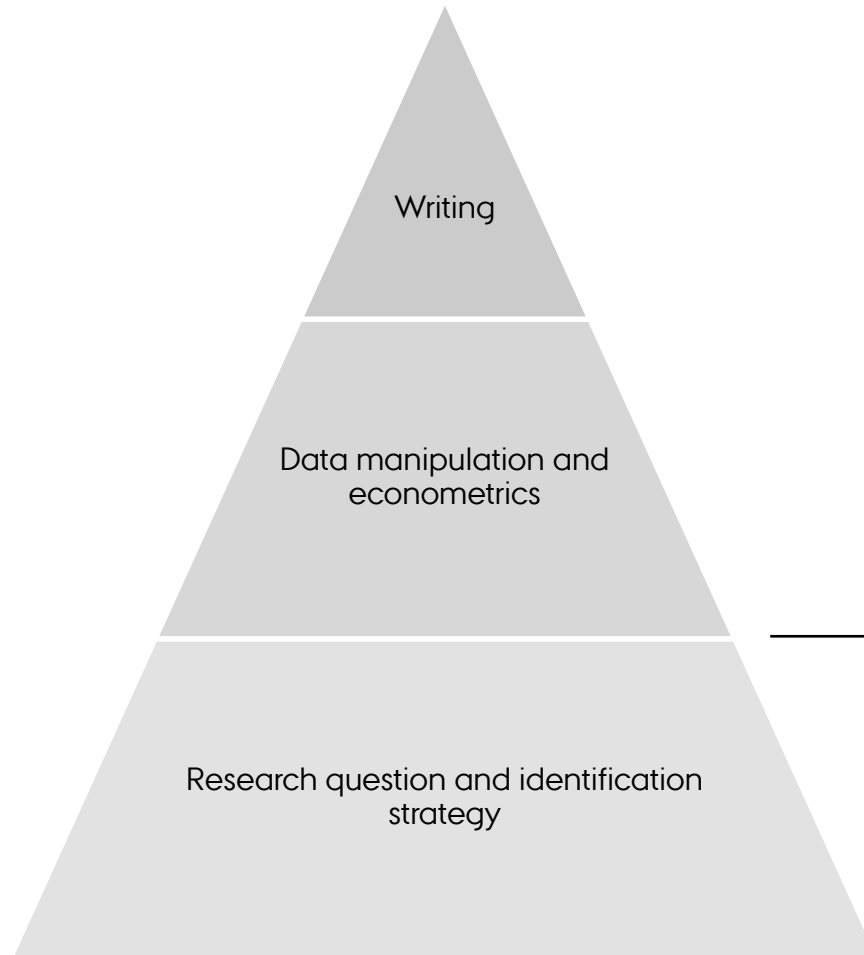- Is the setup allowing causal interpretation?

Pyramid labels (top to bottom): Writing; Data manipulation and econometrics; Research question and identification strategy

# Agenda

Data sources and research strategies

**Research results (Administrative, commercial, and registry data)**

Perspectives ahead

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Prof. Dr. Steffen Meyer

# Example 1 - Administrative:
# Does uncertainty in financial markets drive investment decisions?

Writing

Data manipulation and econometrics

Research question and identification strategy

**Key questions:**

- Administrative data used in numerous research papers
- Administrative data enhanced by (unmatched) survey data
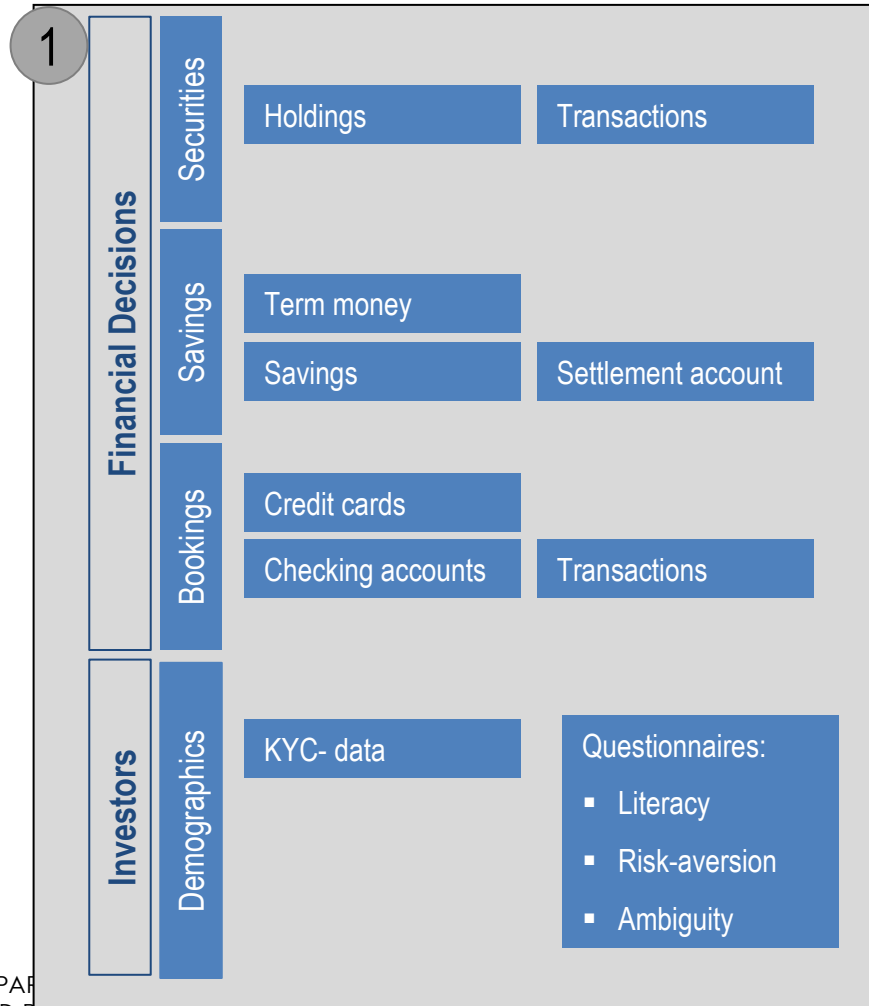- Simple regression designs with clustering

**Key questions:**

- No causal evidence on the effect of uncertainty on equity investments
- Equity participation relevant for practice, society and research
- We use fund liquidations for identification

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Prof. Dr. Steffen Meyer

# Charline Uhr and I, combine the fund liquidations with administrative data on client trades and measures of ambiguity

**Databases of investor data combined with (100k investors, 1999 to 2006)**

**(1)** **Financial Decisions**

**Securities**
- Holdings
- Transactions

**Savings**
- Term money
- Savings
- Settlement account

**Bookings**
- Credit cards
- Checking accounts
- Transactions

**Investors**

**Demographics**
- KYC- data
- Questionnaires:
  - Literacy
  - Risk-aversion
  - Ambiguity

**Mutual fund terminations (since 2006)**

**(2)**
- 2,222 fund terminations (obtained from Bundesverband Investment und Asset Management e.V.) for 1,958 distinct clients

**Measures of ambiguity**

**(3)**
- Market-based
  - Volatility-of-volatility measures (VVIX (V-VSTOXX), 30-day forward price of the VIX (VSTOXX))
  - Omega following Brenner & Izhakian (2018)
- Expert-survey-based
  - Dispersion of opinion of professional forecasters

Prof. Dr. Steffen Meyer

DEPAR... AND BUSINESS ECONOMICS AARHUS UNIVERSITY

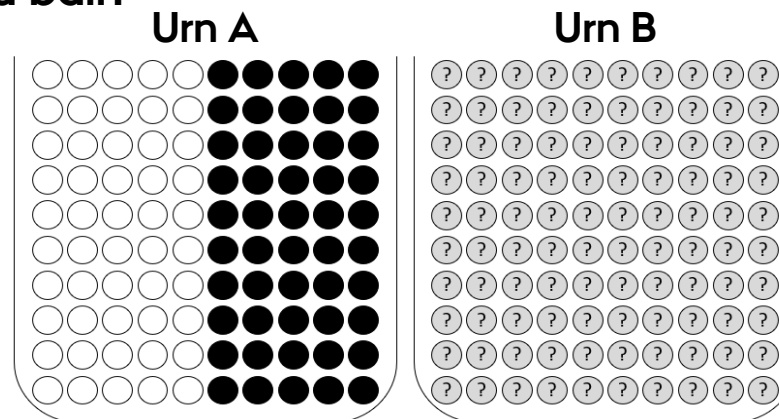# Does ambiguity influence the decision of people? The Ellsberg (1961) experiment

**The Ellsberg experiment:**

- You can draw one ball from urn A or urn B.

- If the ball is black, you win 10,000 Euro.

*Your choice is between Urn A and Urn B:*

- In Urn A, there are 30 white and 30 black balls.

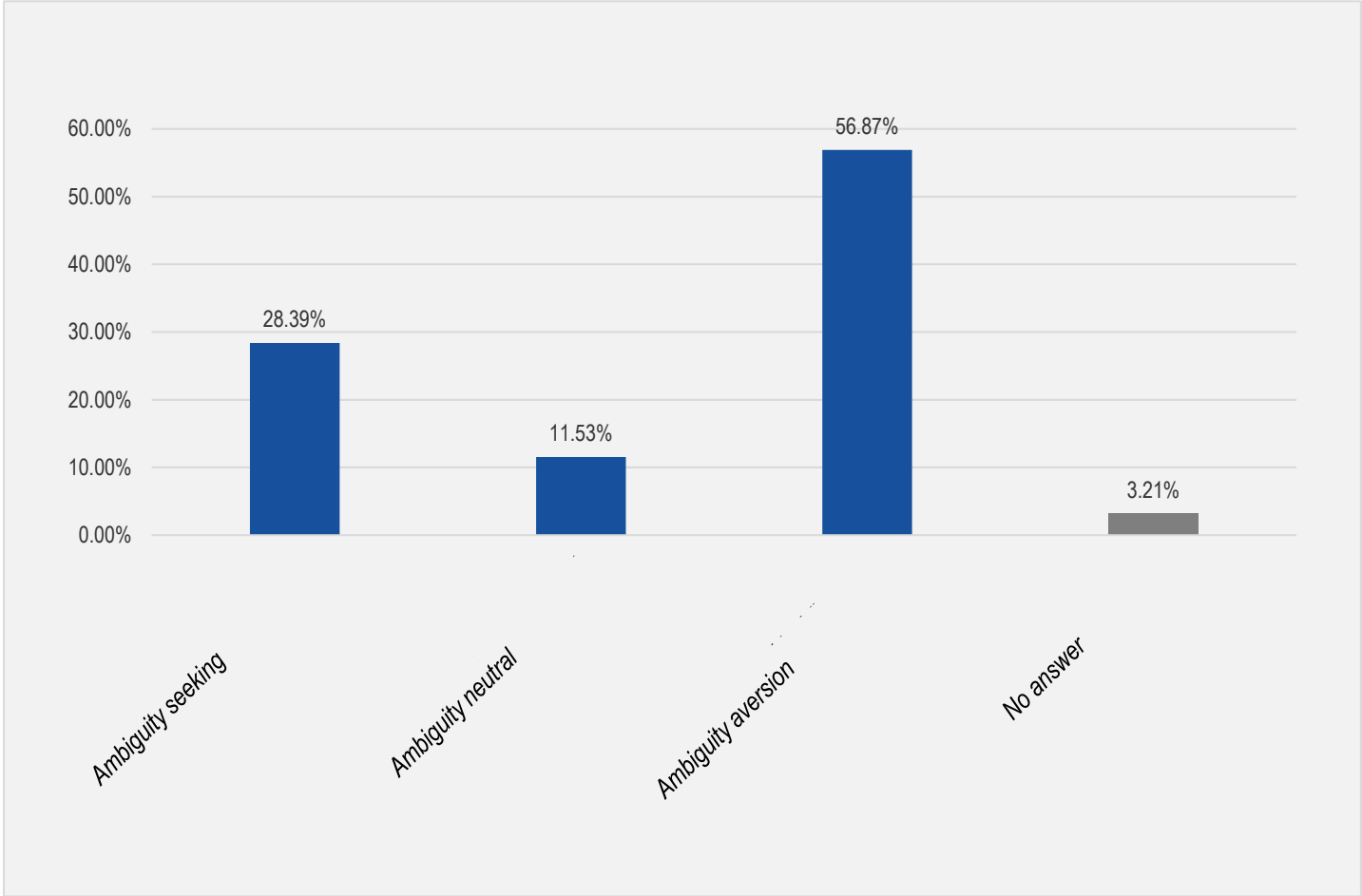- In Urn B, there are 60 balls. They can be either black or white.

→ **Which urn do you pick to draw a ball?**

**Urn A**          **Urn B**



Source: Kostopolous, Meyer, Uhr (2022): Ambiguity about volatility and investor behavior, Journal of Financial Economics.

# Playing the thought experiment in the real world, shows that 60% of people are ambiguity averse and pick urn A
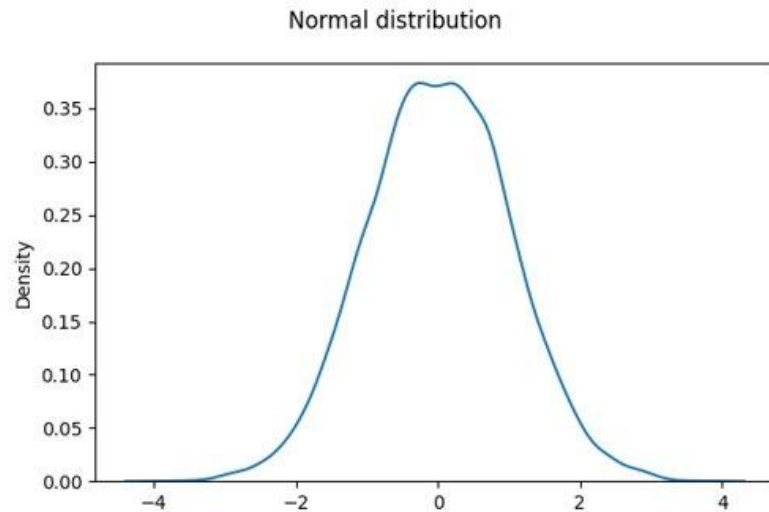
The (REAL) experiment (N = 997):

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Prof. Dr. Steffen Meyer

# Ambiguity is distinct from risk in that an agent cannot assign unique probabilities to all events
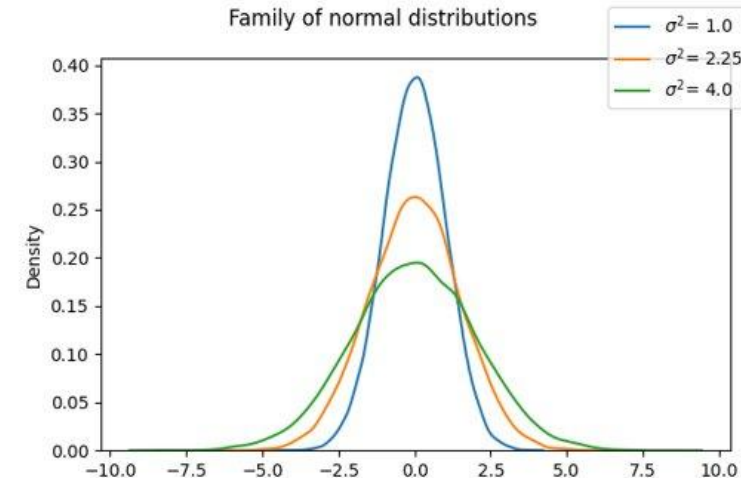
## Risk regarding future outcomes

- Future outcomes are uncertain but can be associated with a probability distribution
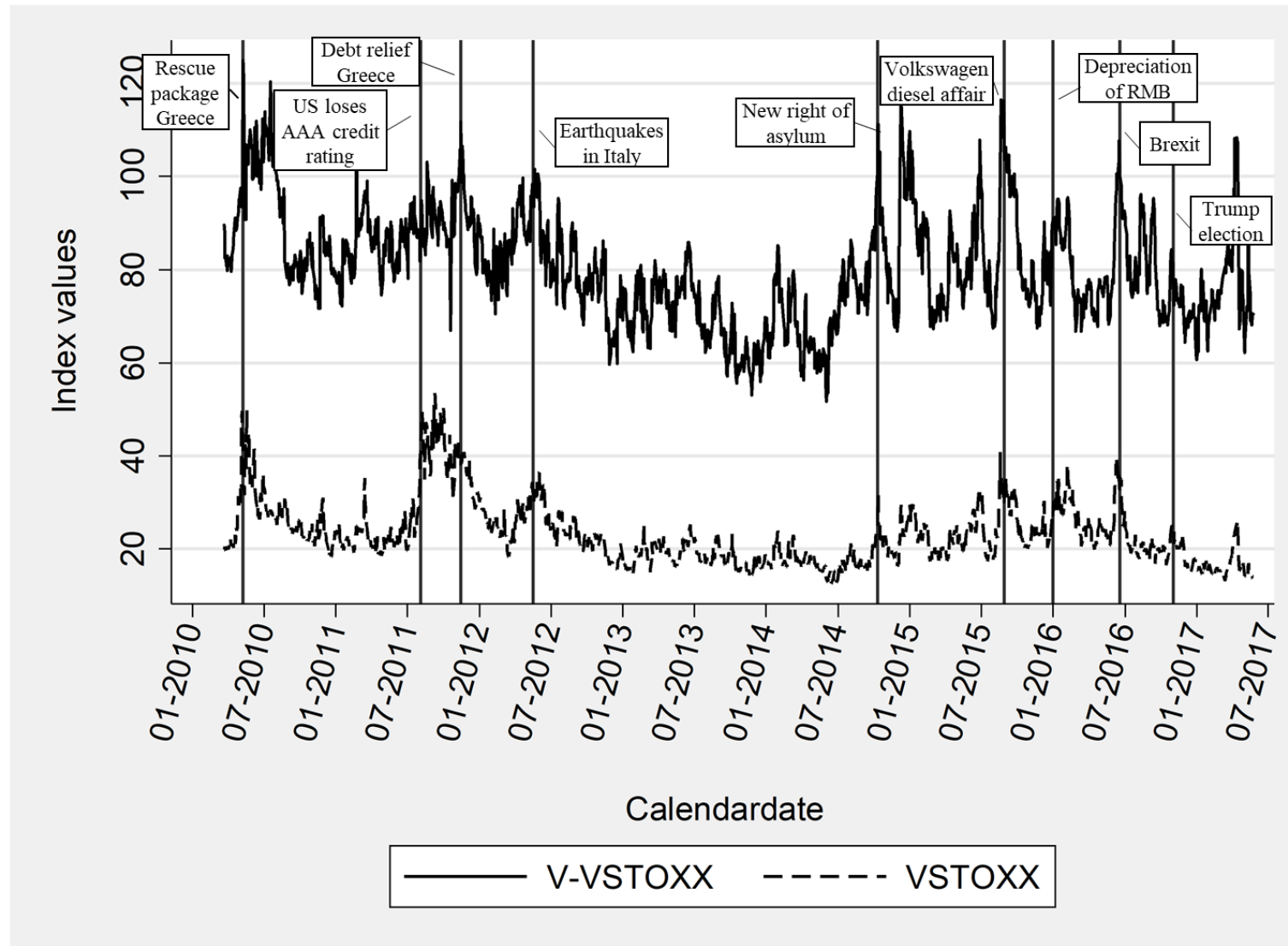- The probability may be **subjective** but is **unique** from the agent's point of view



## Ambiguity about volatility

- The standard deviation is ambiguous, tracing out several distributions
- Each outcome can be associated with **several** probabilities

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

AARHUS BSS

Prof. Dr. Steffen Meyer

AACSB ACCREDITED   AMBA ASSOCIATION ACCREDITED   EQUIS EFMD ACCREDITED

10

# „Eyeballing" suggests that V-VSTOXX (also VVIX) makes sense as one sensible ambiguity measure

Prof. Dr. Steffen Meyer

# To study the effect of ambiguity, we use fund terminations which allows isolating the effect of ambiguity
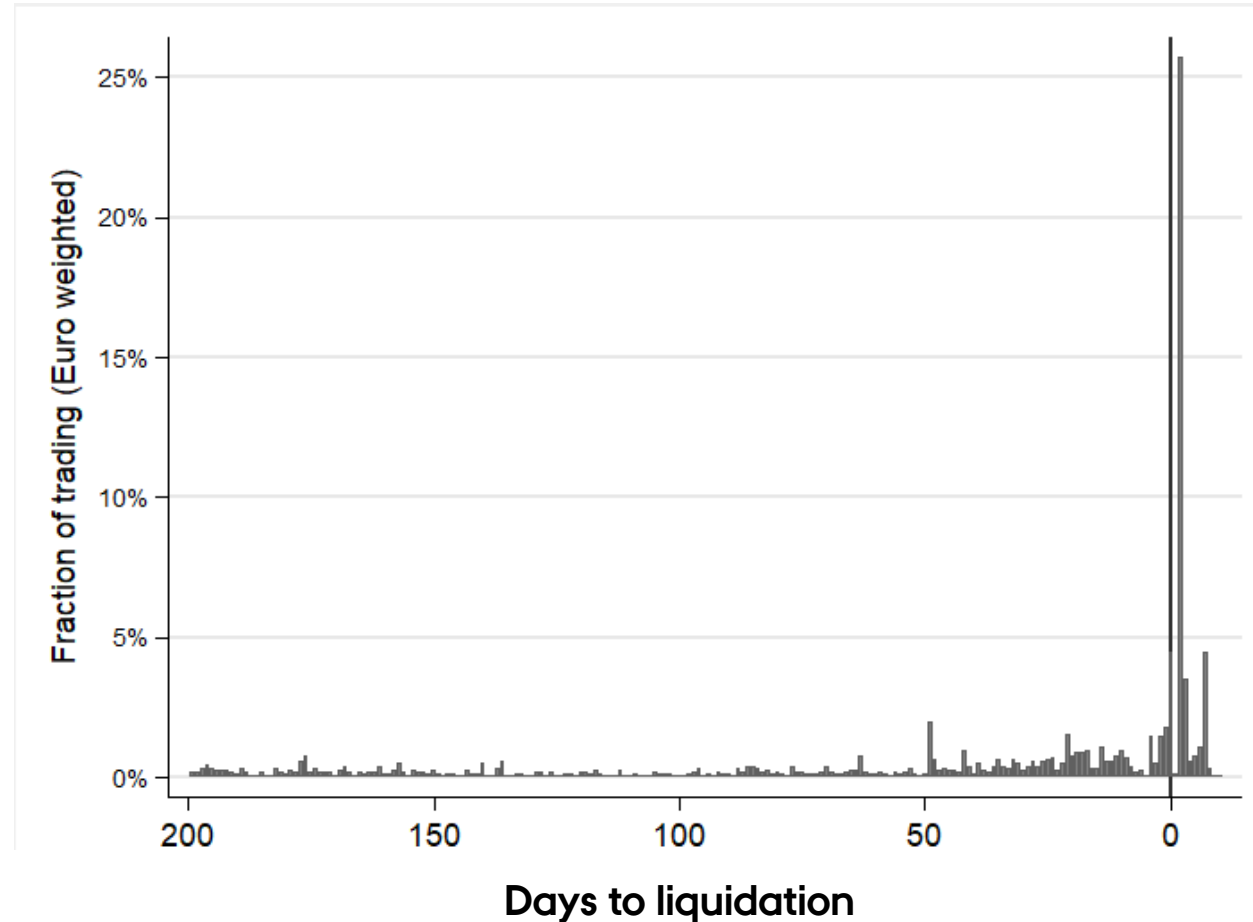
Advantages of using fund terminations:

1.  Fund terminations although announced, hit private investors by surprise

2.  Fund terminations neither affect the wealth nor the gains and losses of investors

3.  No endogeneity from liquidity needs, treatments or windfalls

4.  Refund is a salient and sizable event to clients and prompts a re-investment decision

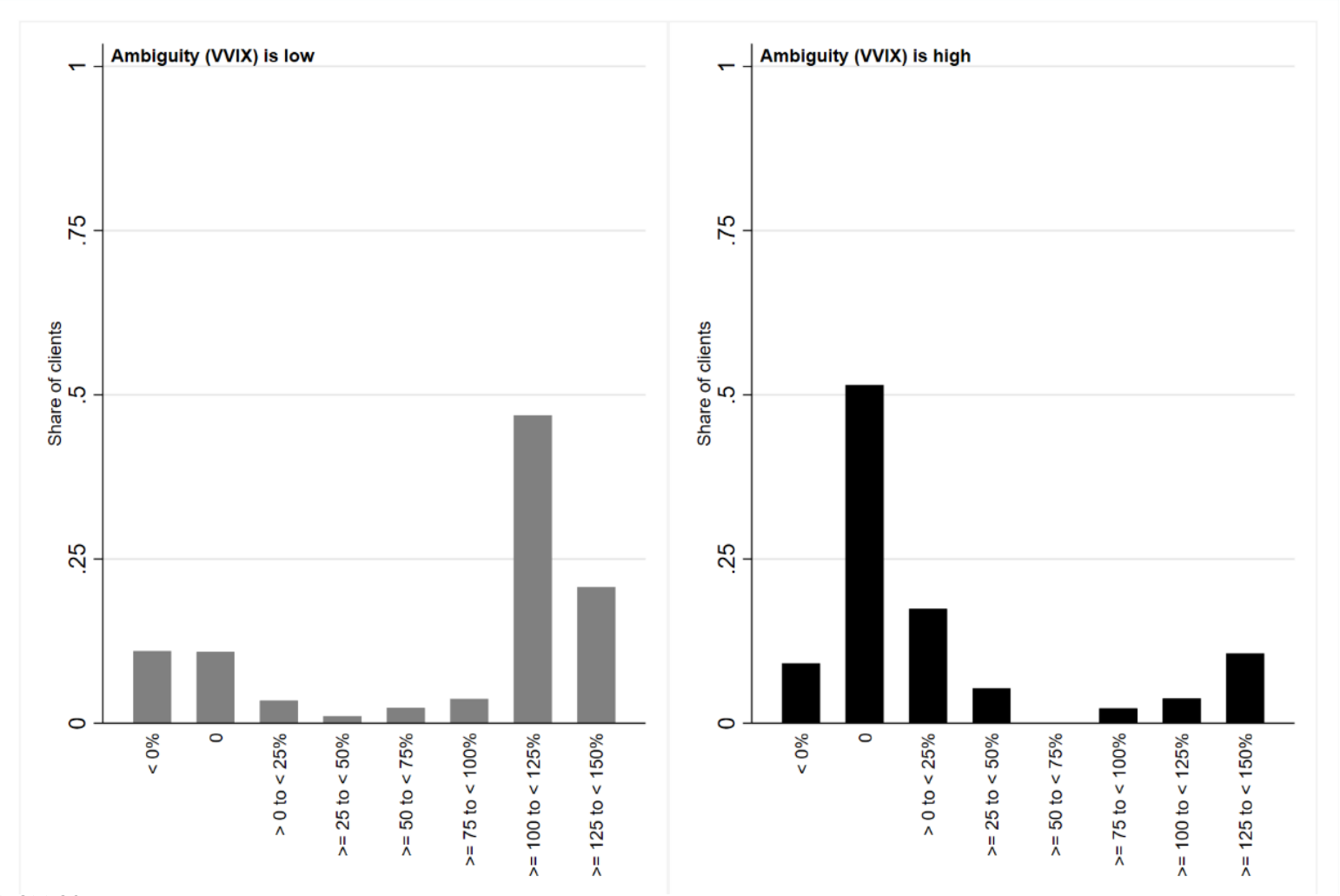5.  Ambiguity at the fund termination date cannot be managed or foreseen by the fund management firm

# Fund terminations must be announced publicly 6 months to 4 weeks before termination. No private communication to investors by fund or bank, so the liquidation hits them as a surprise

Trading behavior before and after fund termination:

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

# One paper in one graph: The level of ambiguity significantly alters investment decisions by private investors
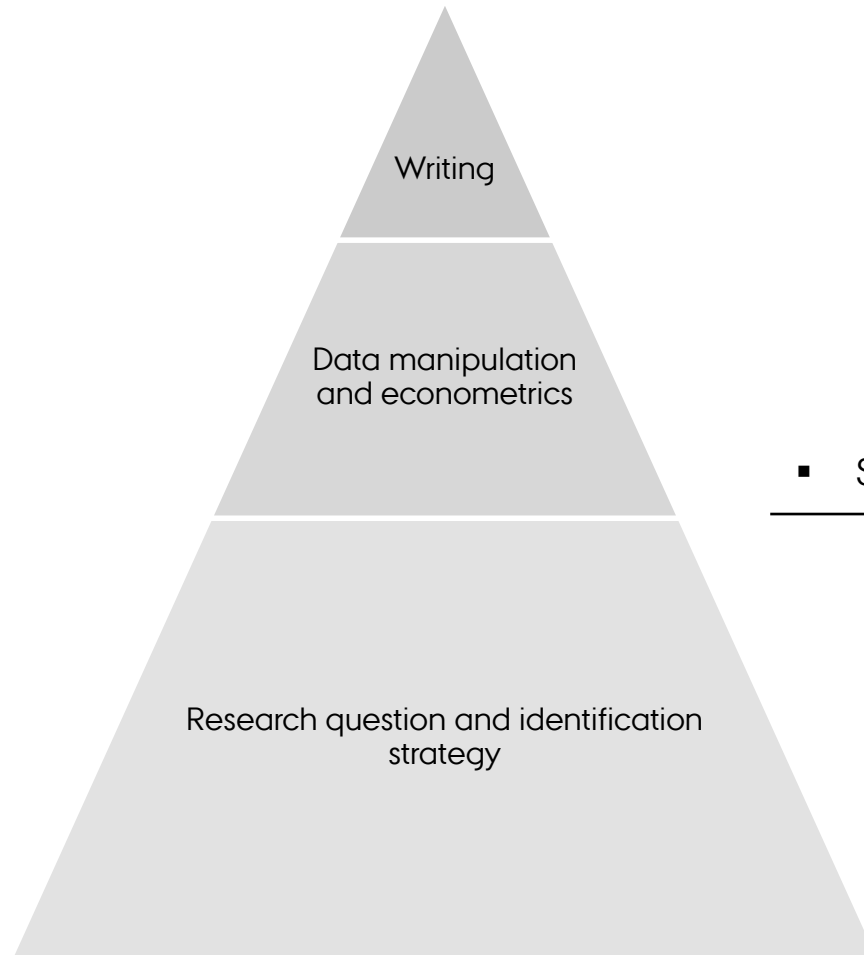
Reinvestment decisions of 2,222 investors after fund terminations:

# Example 2 – Commercial data:
# <u>Does</u> uncertainty affect the incorporation of information in financial markets?

Writing

Data manipulation and econometrics

Research question and identification strategy

**Key questions:**

- Data from Datstream

- Codes on Github

- Simple regression designs with clustering, except for long-term convergence

**Key questions:**

- No causal evidence on the effect of uncertainty on information processing

- The incorporation of information on capital markets is a key function to the financial system

- Exploit time-varying ambiguity and pre-defined earnings announcement dates

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

AARHUS
BSS

Prof. Dr. Steffen Meyer

AACSB ACCREDITED  AMBA ASSOCIATION ACCREDITED  EFMD EQUIS ACCREDITED

15

# Theory on information processing under ambiguity: Uncertainty about volatility induces an asymmetric response to good (bad) news signals
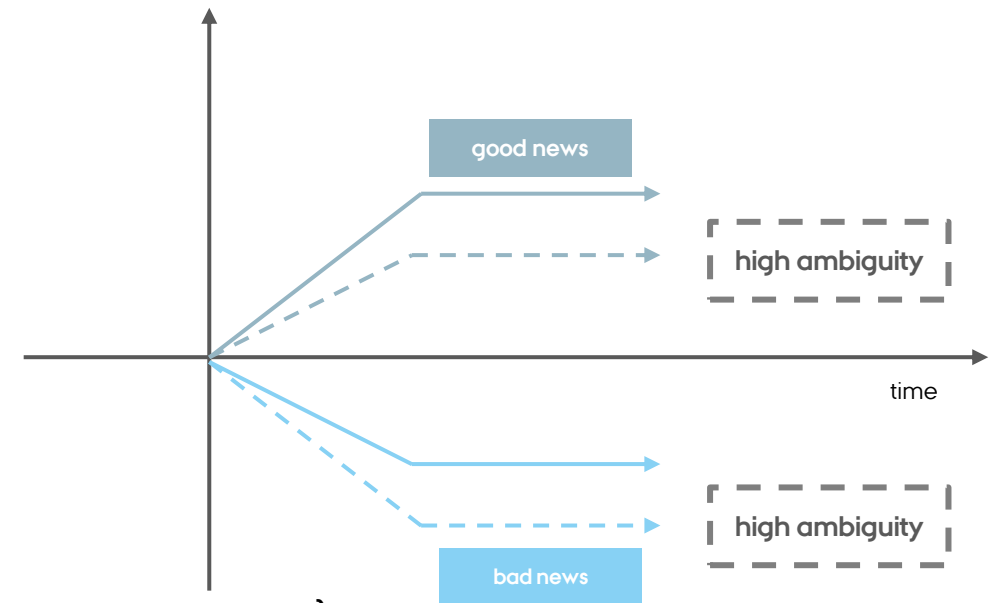
Ambiguity, Information Quality, and Asset Pricing, Epstein & Schneider (2008)

## The set-up

- Agents evaluate assets based on their market covariance and a noisy signal
- The signal noise is ambiguous, i.e. true standard deviation is uncertain
- The agent imputes the uncertainty through a family of distributions (i.e., varying sigma)
- Agents evaluate utility in a min-max framework

## The implication

- The worst/ best belief about credibility differs between positive and negative surprises
- Positive surprises reach their worst case when they are least credible (=> Underreaction)
- Negative surprises reach their worst case when seen as most credible (=> Overreaction)

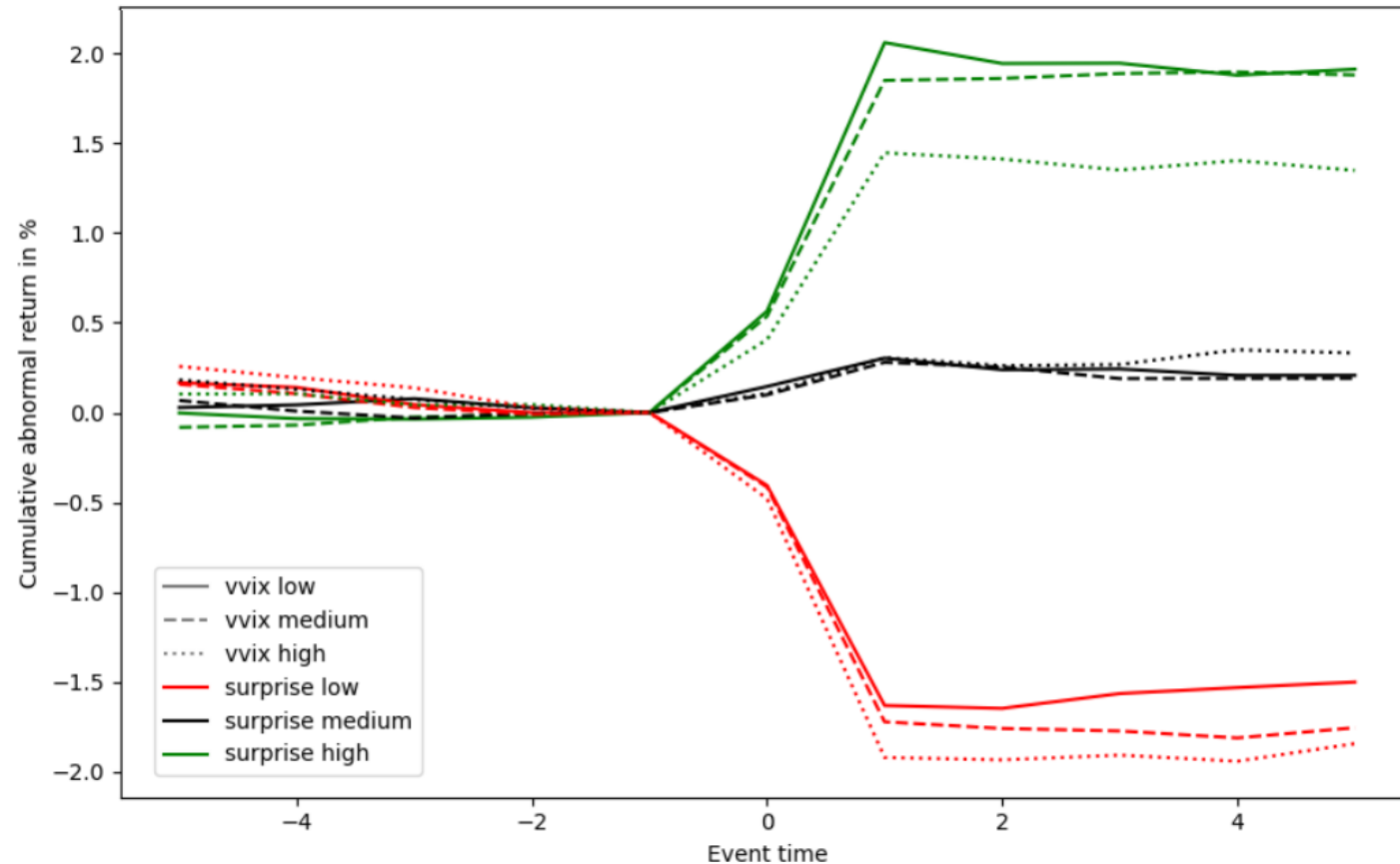# We use over 30'000 earnings announcements in a sample of 616 S&P 500 firms spanning from 2006 to 2023

**Our sample**

▪ Survivorship free (annual) sample of the S&P 500 spanning 2006 to 2023, obtained from Datastream and Eikon

▪ The sample encompasses 34'039 earnings announcement events, i.e. 55 events per firm on average
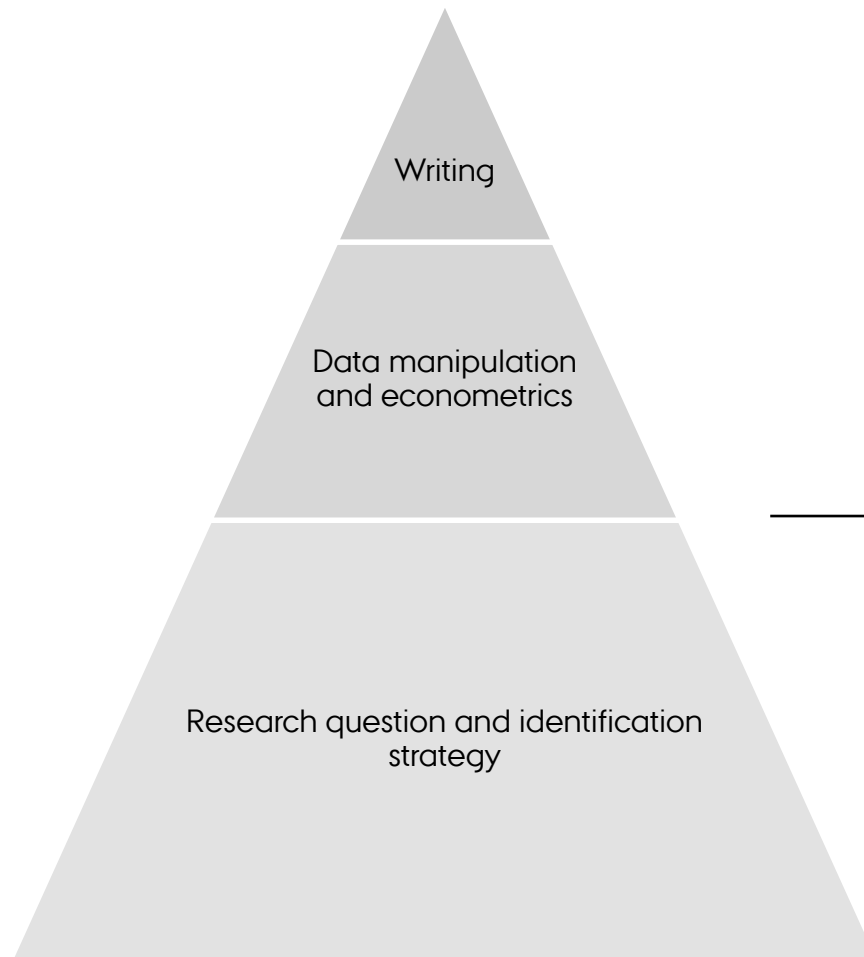
| Variable | Proxy | Frequency | Source |
|---|---|---|---|
| Ambiguity | Volatility of expected volatility on S&P 500 options (VVIX) | Daily | CBOE |
| | Variance of growth probability forecasts from Survey of Profession Forecasters (SPF) | Quarterly | Philadelphia FED (own calculation) |
| | Economic Policy Uncertainty Index (Baker, 2016) | Monthly | Scott Baker |
| | Volatility of return probabilities (Omega) | Weekly | Datastream (own calculation) |
| Market value | Firm size | Daily | Datastream |
| Market to Book ratio | Firm valuation | Quarterly | Datastream, Ikon (own calculation) |
| Idiosyncratic risk | Average squared residual of Fama-French 3 factor model | Daily (annual moving window) | Datastream (own calculation) |
| Short-sales constrains | Share of institutional ownership | Bi-annual to annual depending on firm | Eikon |
| Firm-level market disagreement | Dispersion of analyst EPS expectation | Quarterly | Eikon |
| Leverage | Firm leverage | Quarterly to annually | Eikon |
| Firm market risk exposure (beta) | Fama-French market factor estimate | Daily (annual moving window) | Datastream (own calculation) |
| Momentum returns | Exposure to momentum factor | Daily (annual moving window) | Datastream (own calculation) |

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

AARHUS
BSS

Prof. Dr. Steffen Meyer

17

# When aggregating over ambiguity and surprises, we observe an asymmetric response in cumulative abnormal returns, as suggested by theory

Market reactions to earnings announcements for different levels of ambiguity

Prof. Dr. Steffen Meyer

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

# Example 3 – Registry data (Nordics): Does more information on pension (gaps) alter investment decisions?

Writing

Data manipulation and econometrics

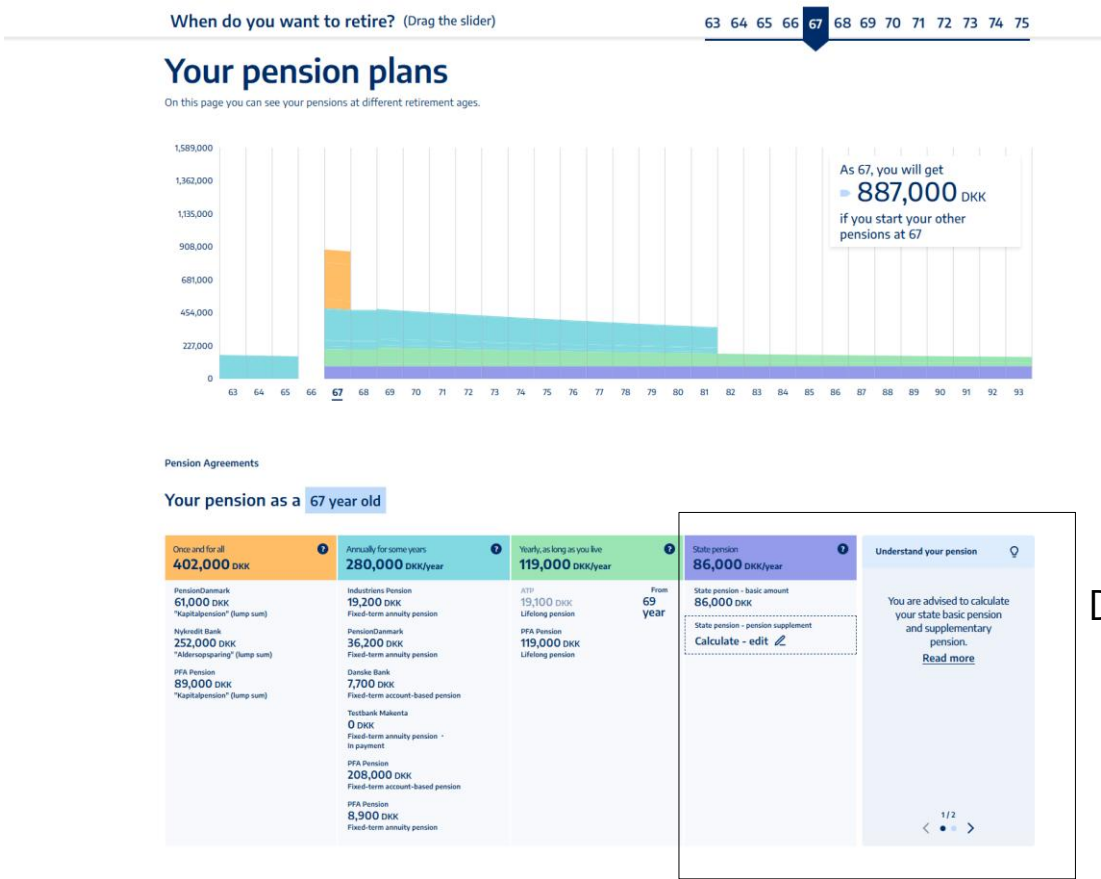Research question and identification strategy

**Key questions:**

- Danish registry data
- Codes on the servers of Statistics Denmark
- Simple regression designs with clustering for a diff-in-diff setting

**Key questions:**

- No causal evidence on the effect of providing information on pensions on the subsequent investment behavior of households
- Over- or undersavings may have direct effects on consumption, growth, and potentially old-age poverty
- Exploit staggered introduction of "pensionsinfo.dk" for different pension plans

# Pensionsinfo.dk was rolled out starting in 2008, leaving out university employees who only had basic access

Pension information from pensionsinfo.dk



Default information

# The setup allows measuring the effects of information provision to the society

## Setup and first results

**Sample:**
- All Danish households 2004 – 2024

**Content:**
- Logins to pensionsinfo by year
- Investments
- Housing
- Education
- Health
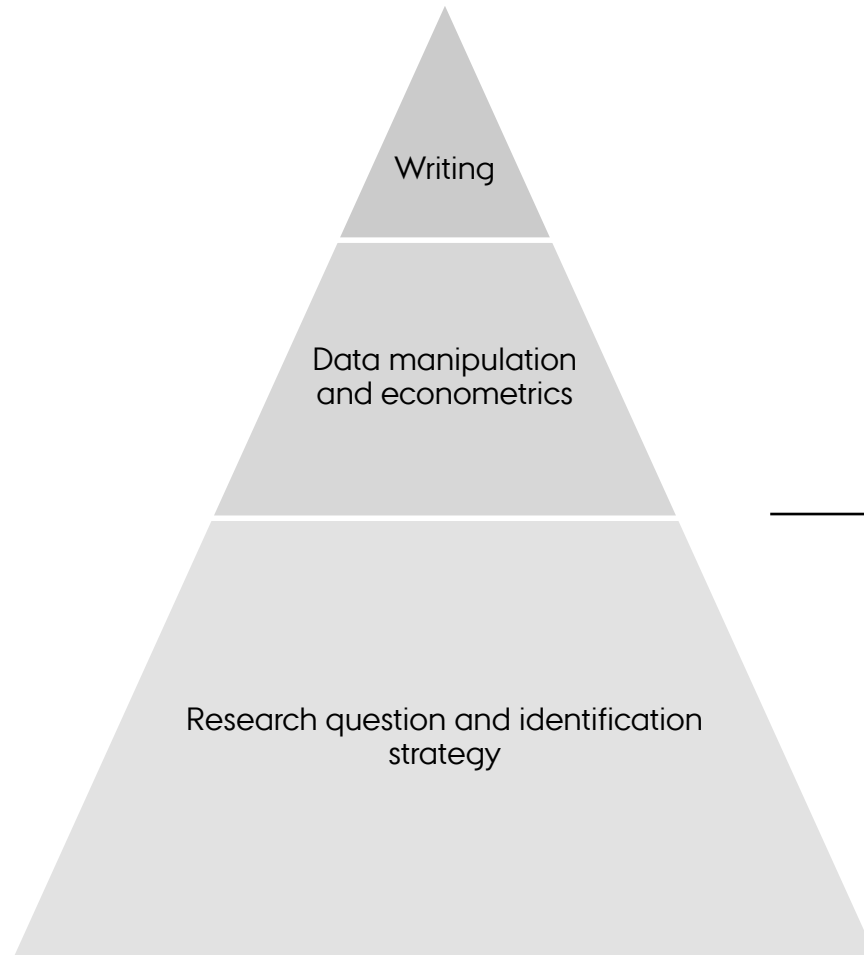- Family situation

**Frequency:**
- Annual

■ Result 1: The ones who need it more, use it less
- ■ Less educated
- ■ Less wealthy
- ■ Younger

■ Result 2: Logins and actions correlate
- ■ Logging in and smarter decisions seem correlated
- ■ But: also, logins by the control group seem to trigger action
- → Is more information really causing action?

Prof. Dr. Steffen Meyer

# Example 3 – Registry data (Germany): Is the 'Wolf of Wall Street' a threat to private investors?

Writing

Data manipulation and econometrics

Research question and identification strategy

**Key questions:**

- Administrative data
- BAFIN "official" data
- Simple regression designs with clustering for a diff-in-diff setting

**Key questions:**

- Is stock spam affecting private investors?
- Negative experiences bear negatively on stock market participation and old-age savings
- Using the time variation in spam campaigns and trading in tout stocks at the same time

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Prof. Dr. Steffen Meyer

# Do you remember the movie „Wolf of Wall Street" and Aerotyne International in it?



**Jordan Belfort:**

"**Aerotyne Industries** is a **cutting-edge high-tech firm out** of the Midwest awaiting **imminent patent approval on the next generation of radar detectors** that have **both huge military and civilian applications** now. […] Your profit on a mere **$6,000** investment would be upwards of **$60,000!**"

# Today such scams are still around and are distributed using emails. Here is the the example of "Equipment & Systems Engineering, Inc." circulating on Dec 1st.

**E-Mail of Dec 1st, 2006:**

Alert! Watch this one Rise!

Equipment & Systems Engineering, Inc. (EQSE)

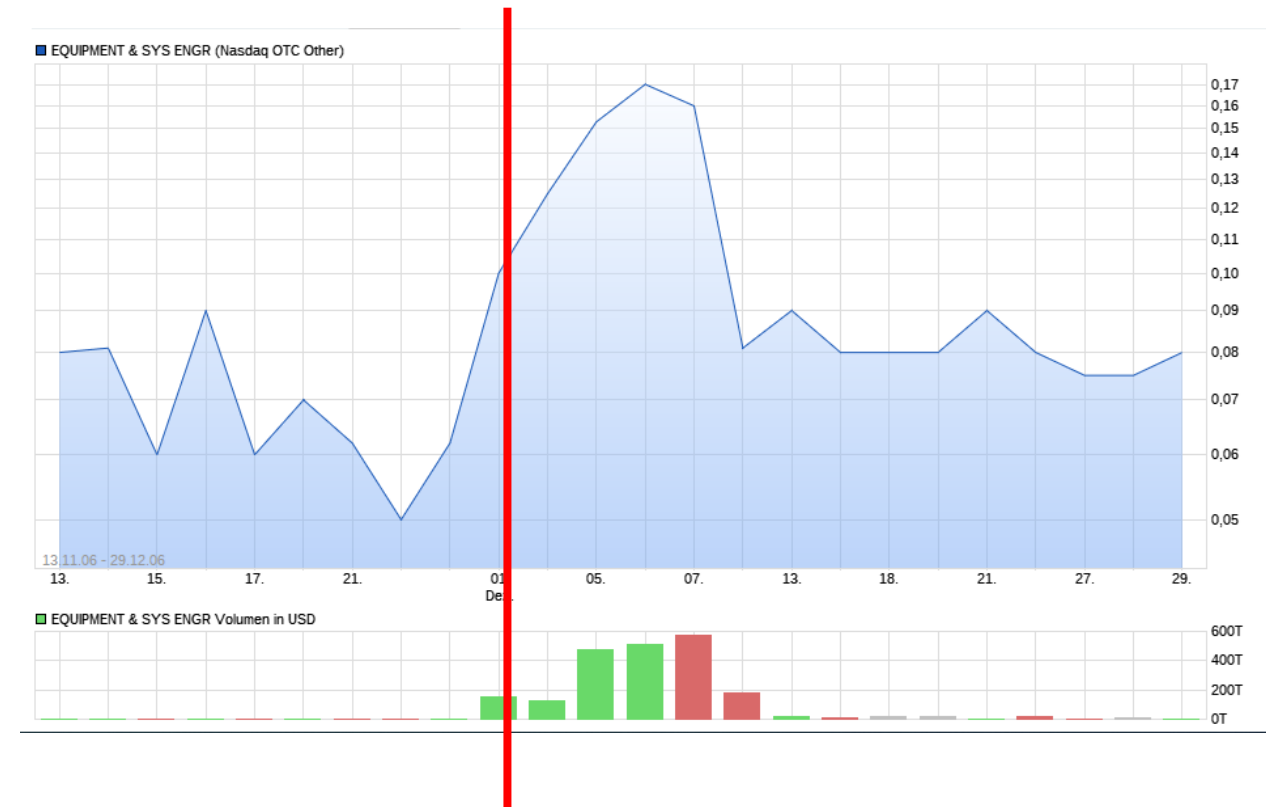Sector: Environmental Engineering
Rating: VERY Bullish
Current Price: $0.062
Projected: $0.20

This is THE pick for the fourth quarter. Environmental stocks are getting incredible exposure and taking off as governments and companies realize they need to start investing NOW. EQSE is involved in high tech solutions for both the problems we face today and those of the future.

Search your favorite financial information site and become a believer. This is HOT!

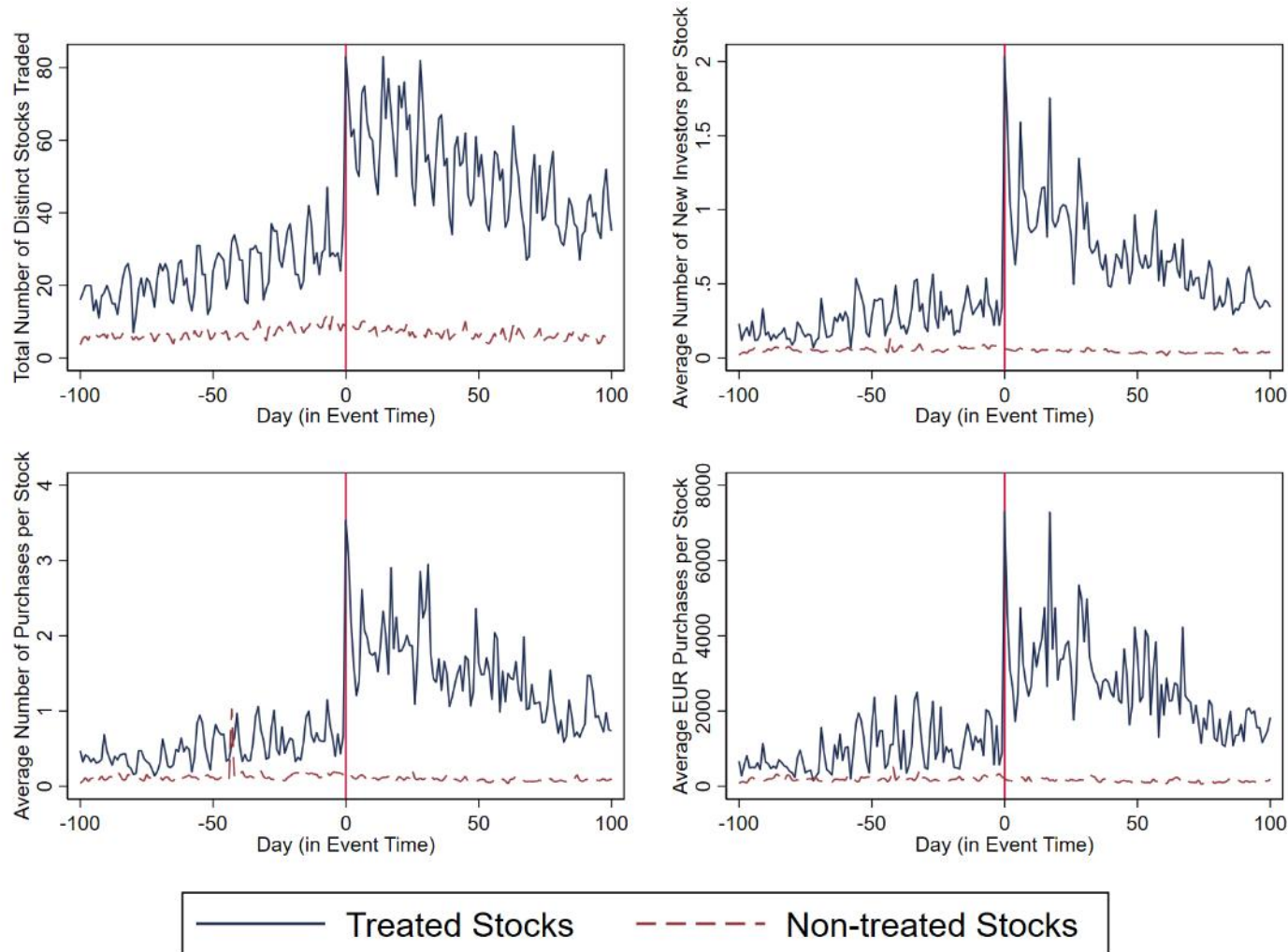An earth-shattering release is expected out of the company any day. With all of EQSE's governmental contacts we are expecting a major contract announcement. This issue is VERY tightly held and the release is going to push it up rapidly. Don't delay. It's not going to stop till we see 20 cents!

**Chart of ESQE in November / December 2006:**

DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Prof. Dr. Steffen Meyer

# Surprisingly people react to the fraudulent emails and obviously believe they could beat the market

Trading volume in treated and (similar) non-treated stocks



**Data:**
- 10 years
- 100.000 investors
- 470 pump-and-dumps
- +/- 100 days around the „event'

**Result:**
- 80% of touted stocks receive a trade
- 30.000 trades in total
- 8.500 investors
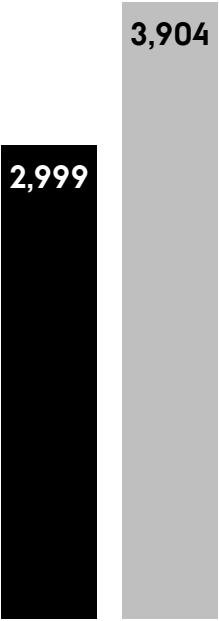
Quelle: Leuz, Meyer et al. (2024)

# As a result, people lose significant amounts of money when investing in such schemes

Performance of investors in pump-and-dump stocks compared to other stock investments



**Number of Trades**

- 5 — Pump-and-dump
- 31 — Normal

**Investment**

- 2,999 — Pump-and-dump
- 3,904 — Normal

**Loss per Trade**

- 31% — Pump-and-dump
- 4% — Normal

Quelle: Leuz, Meyer et al. (2024)

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

AARHUS
BSS

Prof. Dr. Steffen Meyer
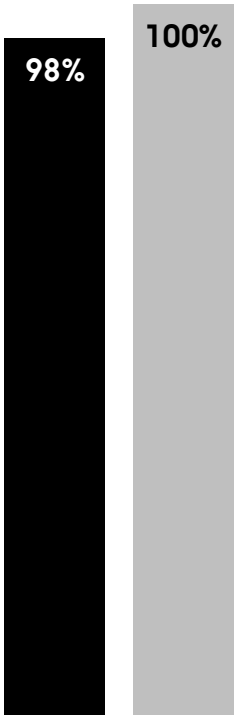
26

# The negative experiences have long-term ramifications on the investments of affected customers

Affected vs. not-affected investors after 12 months in %



**Stock investing**

98% (Pump-and-dump) | 100% (Normal)

■ Pump-and-dump ■ Normal

**Blue-Chip stocks**

40% (Pump-and-dump) | 100% (Normal)

■ Pump-and-dump ■ Normal

**Fund investments**

80% (Pump-and-dump) | 100% (Normal)

■ Pump-and-dump ■ Normal

Quelle: Leuz, Meyer et al. (2024)

Prof. Dr. Steffen Meyer

# Our findings, combined with the power of AI, pose potentially a new challenges to investment decisions and the role of regulation and financial advice

- People react to fraudulent emails
- Exposure to fraudulent schemes has negative ramifications on investment decisions

- The cost of producing fraudulent information decreases
- The ‚quality' (= appeal) of fraudulent information increases
    - Better designs
    - Better language

- More spam
- High incident rates
- → Potentially negative effects on trust in stock markets

**Solution?**
- Understand the tool and address its risks
- HI rationality 'judges'

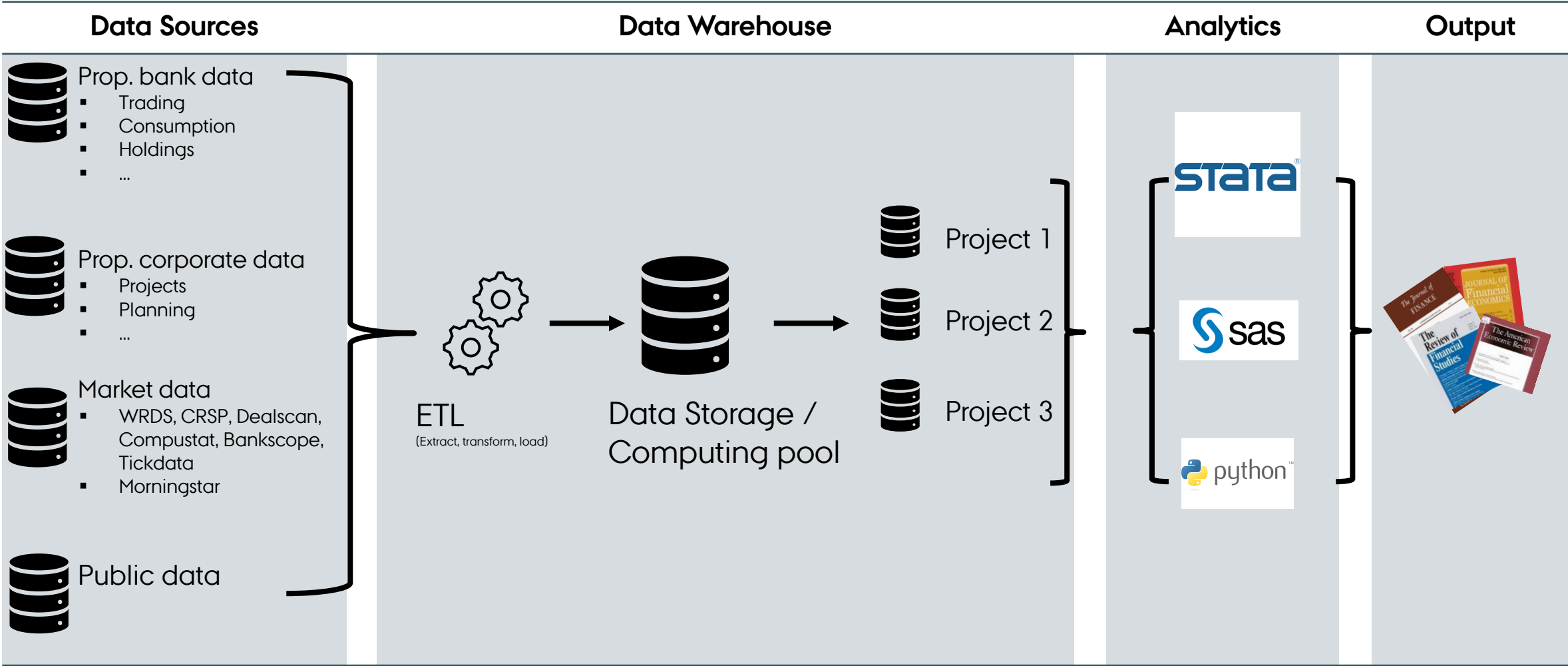DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS AARHUS UNIVERSITY

AARHUS BSS

# Agenda

Data sources and research strategies

Research results (Administrative, commercial, and registry data)

**Perspectives ahead**

# Look ahead: Data-driven research as a trusted partner for the companies is particularly promising in Germany

| Data Sources | Data Warehouse | Analytics | Output |
|---|---|---|---|

**Prop. bank data**
- Trading
- Consumption
- Holdings
- ...

**Prop. corporate data**
- Projects
- Planning
- ...

**Market data**
- WRDS, CRSP, Dealscan, Compustat, Bankscope, Tickdata
- Morningstar

**Public data**

**ETL**
(Extract, transform, load)

**Data Storage / Computing pool**

Project 1

Project 2

Project 3

Prof. Dr. Steffen Meyer

30

# Ambiguity is in our focus: The research agenda at CentR-A at Aarhus University

How does risk, and particularly ambiguity, affect economic decision-making?

### Research stream 1

*Investment behavior*

- Reaction to information provision under ambiguity

### Research stream 2

*Financial decision-making under ambiguity*

- Mortgage refinancing under ambiguity

## Account (trading) data:
Financial decision-making, (More than 100k customers, 24 years of data)
- Stock and fund transactions
- Checking-accounts

## Registry data:
Economic decision-making of the Danish population
- Housing
- Insurance
- Wealth

## Market data:
LSEG Eikon with Datastream, Bloomberg (Factiva), Thomson Reuters Tick History

# "If I had an hour to solve a problem, I'd spend 55 minutes thinking about the problem and 5 minutes thinking about solutions."

*(Albert Einstein)*

DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Prof. Dr. Steffen Meyer

AACSB
ACCREDITED

ASSOCIATION
AMBA
ACCREDITED

EFMD
EQUIS
ACCREDITED