



# Data-Driven Financial Conduct Regulation: *the FCA's remit, datasets and research, and opportunities for collaboration*

Dr Stefan Hunt

Head of Behavioural Economics and Data Science

Big Data Analytics for Financial Services, UCL  
7<sup>th</sup> January 2016

# Remit of the FCA

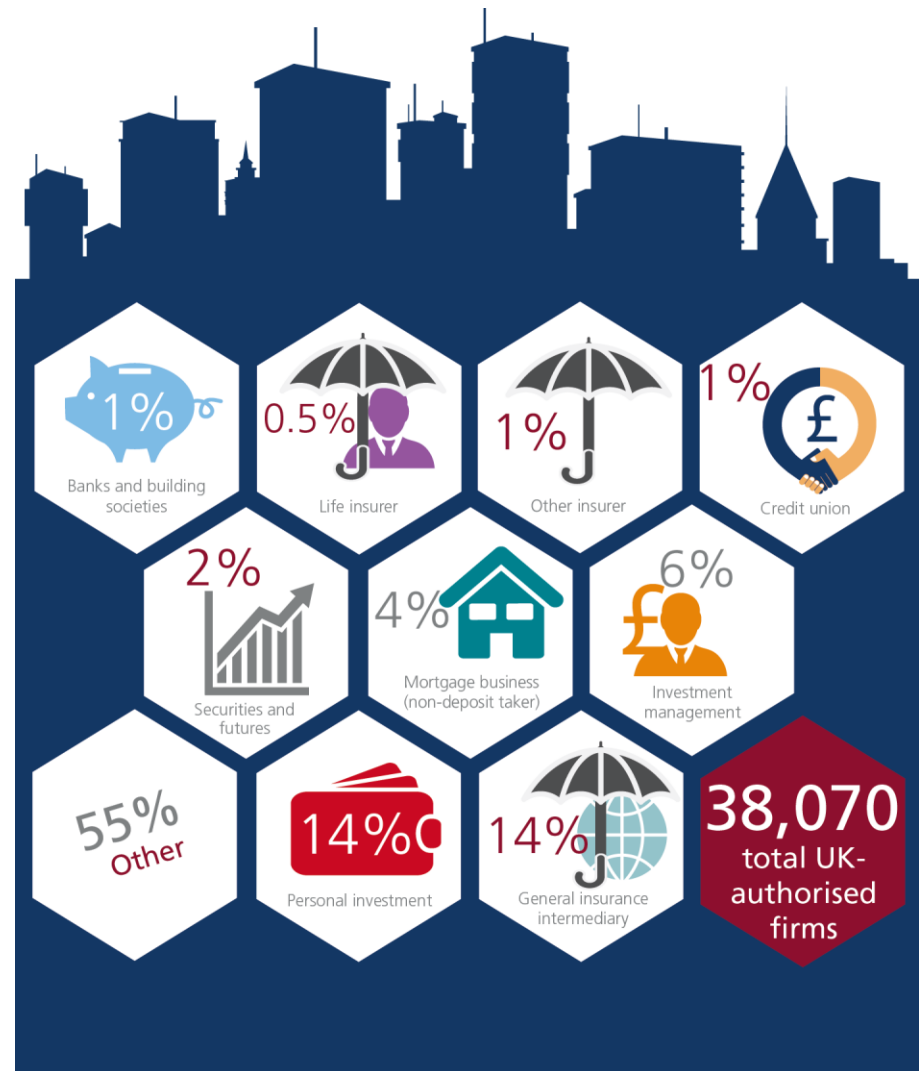
We regulate most of the UK financial markets.

## Retail:

- Savings and investments
- Consumer credit
- Mortgages
- Insurance
- ...

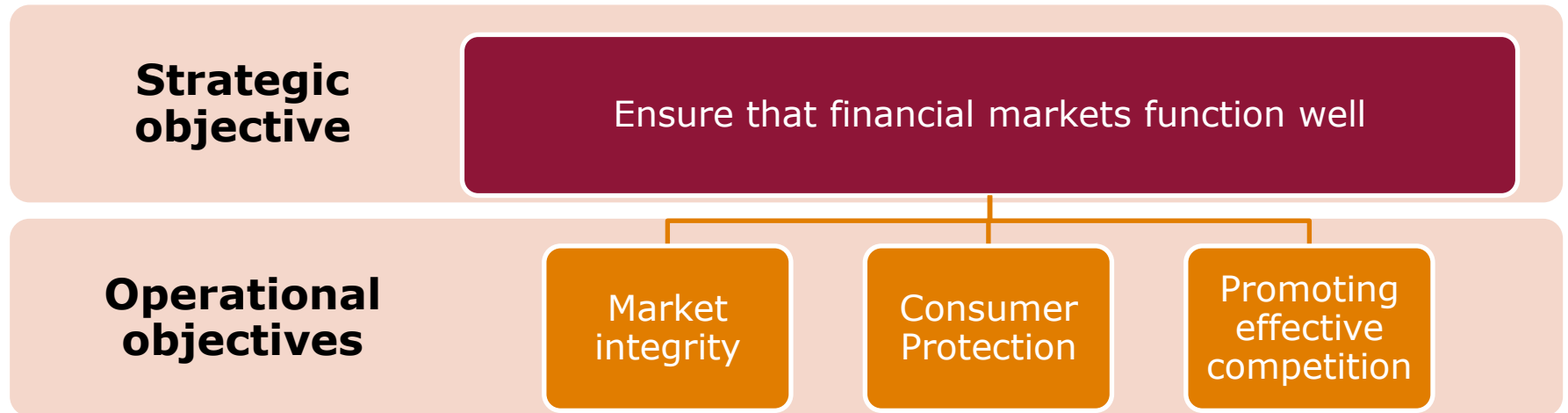
## and wholesale:

- Investment banking
- Fund management
- ...



Number correct as at 6 January 2016. Does not include consumer credit firms with interim permissions. "Other" firms are mainly consumer credit

# Objectives and powers



The FCA intervenes in markets through:

- **Authorising** firms and people to operate
- **Policy-making**: creating laws
- **Supervision**: check compliance
- **Enforcement**: prosecution and punishment

...increasingly using competition analysis

# Key FCA data sets

## *Wholesale:*

- 1. Financial transactions** / Zen
- 2. EMIR** (interest rates, OTC derivatives)
- 3. AIFMD** (hedge funds)

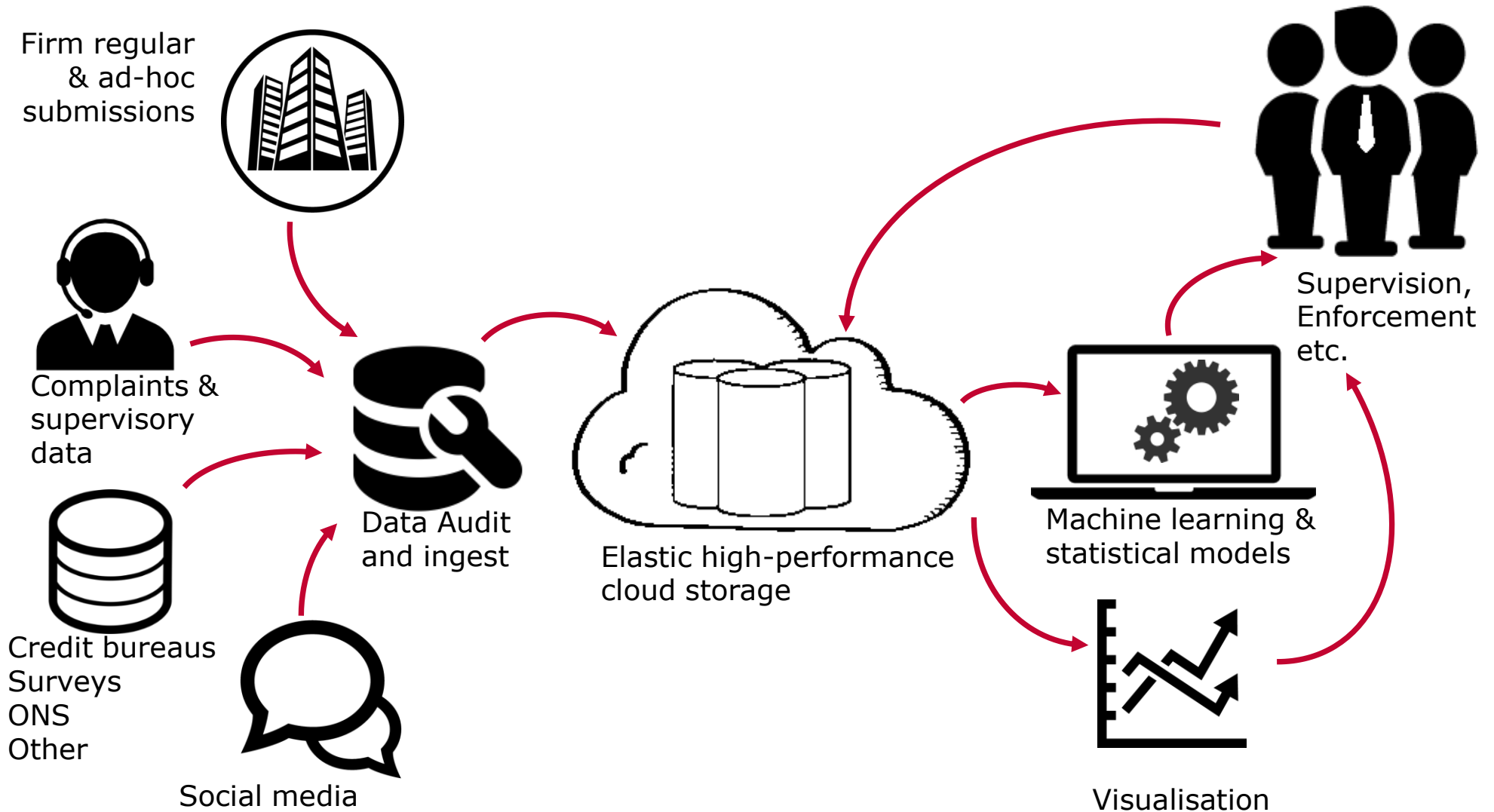
## *Retail:*

- 4. Payday** lending
- 5. Credit card** statements (~ all statements for last five years)
- 6. Credit bureau files**
- 7. Personal current account** micro data
- 8. Data from large field experiments** (e.g. savings, insurance), matched with surveys
- 9. Product sales data** (retail products, mortgages good quality)

## *Firms and employees:*

- 10. Firms'** regulatory submissions, consumer complaints etc.
- 11. Employees'** authorisations and records

# The data ecosystem



# Payday lending price cap

Parliament created duty to impose cap on “high-cost short-term credit”. Structure and level decided by FCA

## Questions:

**1**

What happens to firms and firms' lending decisions?

**2**

What options are there for consumers without access to loans? Are they better or worse off?

# Data

- **Requested data using formal legal powers**
- **Data on payday loans in 2012-3:** top 37 lenders, ~99% market
- **For 11 lenders, ~90% market, all applications,** denied and accepted, including lender credit score and revenues and costs
- **Match applicants across firms and to credit bureau files** using unique identifier. 6 years of data including loan applications, holding and balances, credit events, defaults and credit bureau credit scores
- **Dataset of vast majority of first-time loan applications,** ~1.9million applicants (observe 4.6 million people, ~10% of adult population)

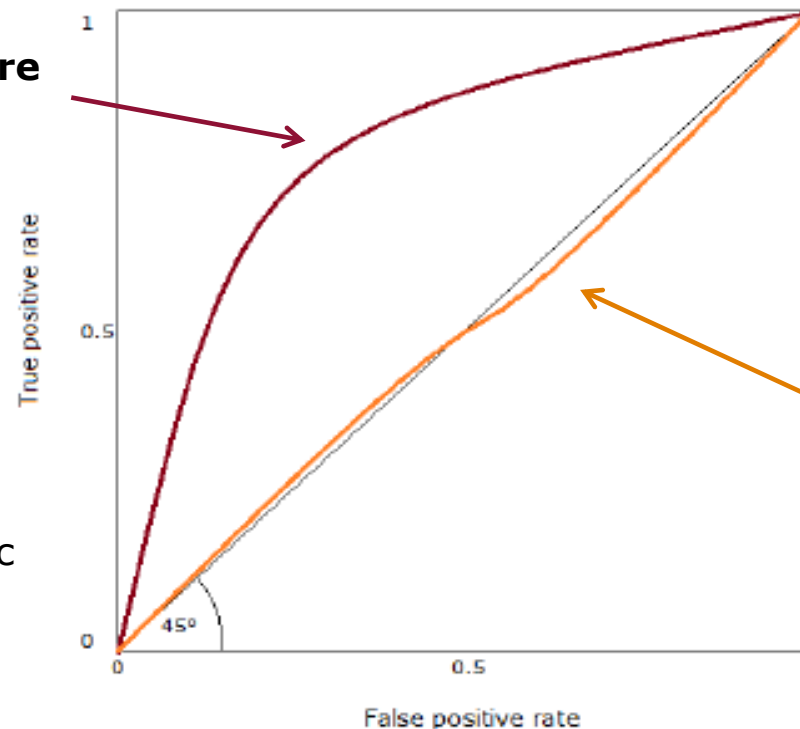
# Recreating lending decisions: credit scores

$$\text{score} = 1000 * (1 - PD)$$

$$PD = \frac{1}{1+e^{-y}}$$

where  $PD$  is probability of default for a given loan.

**'Good' credit score**

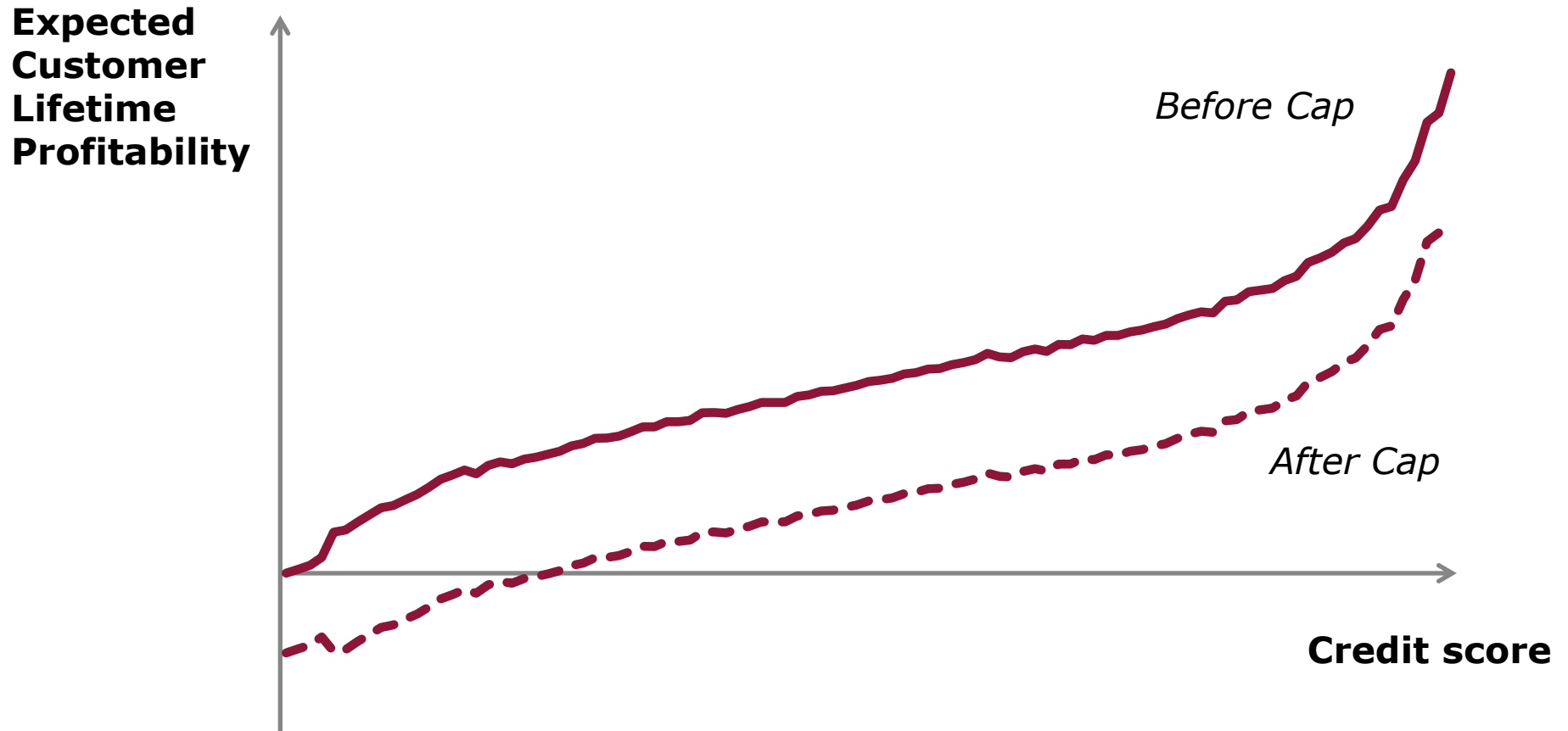


**45° - credit score has no explanatory power**

ROC = Receiver Operating Characteristic



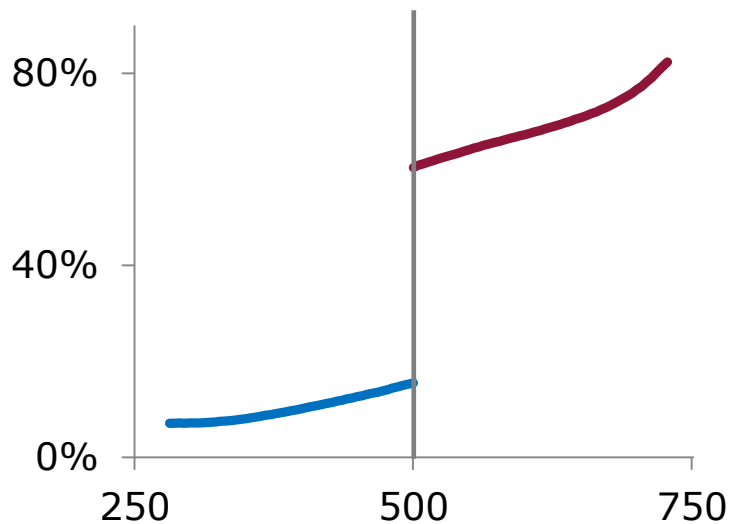
# Example: Impact on customer profitability



# Use regression discontinuity design to estimate causal effect of payday loans

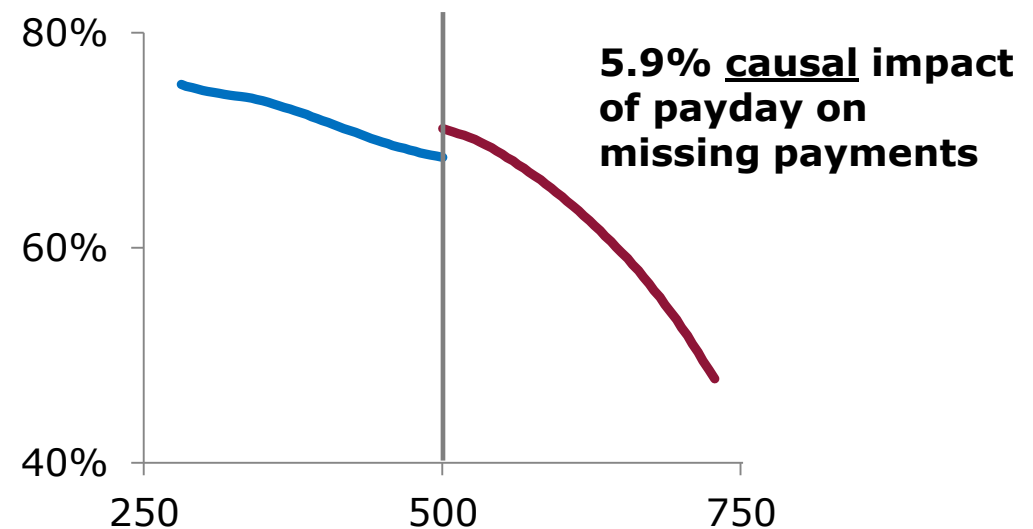
**1<sup>st</sup> Stage:**

**Probability of getting  
payday loan**



**2<sup>nd</sup> Stage:**

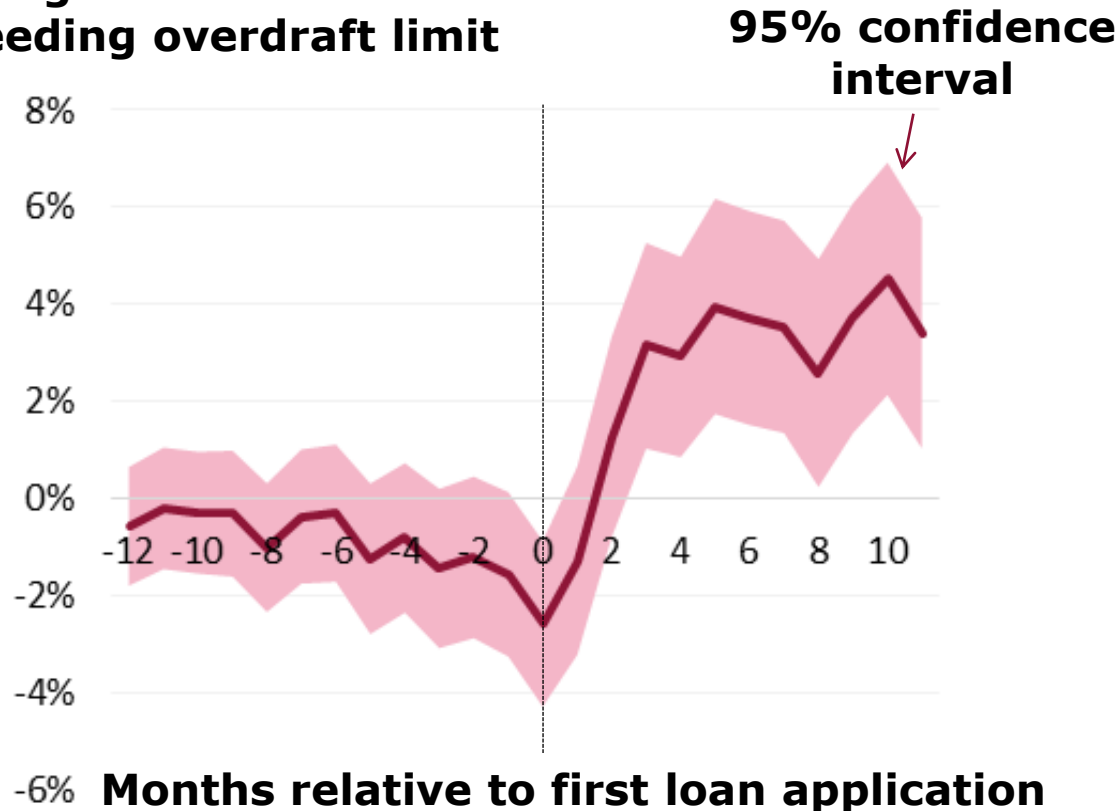
**Probability of missing a non-  
payday payment**



**Internal Credit Score** →

# Causal impact of payday loan use on consumers

Change in likelihood of exceeding overdraft limit



- **Evidence suggests payday use worsens financial outcomes**
- Use behavioural models to assess welfare impacts

**Next step:** identify heterogeneous treatment effects, who is gaining and losing, using data science methods (Athey and Imbens, 2015)

# More practical examples of using research

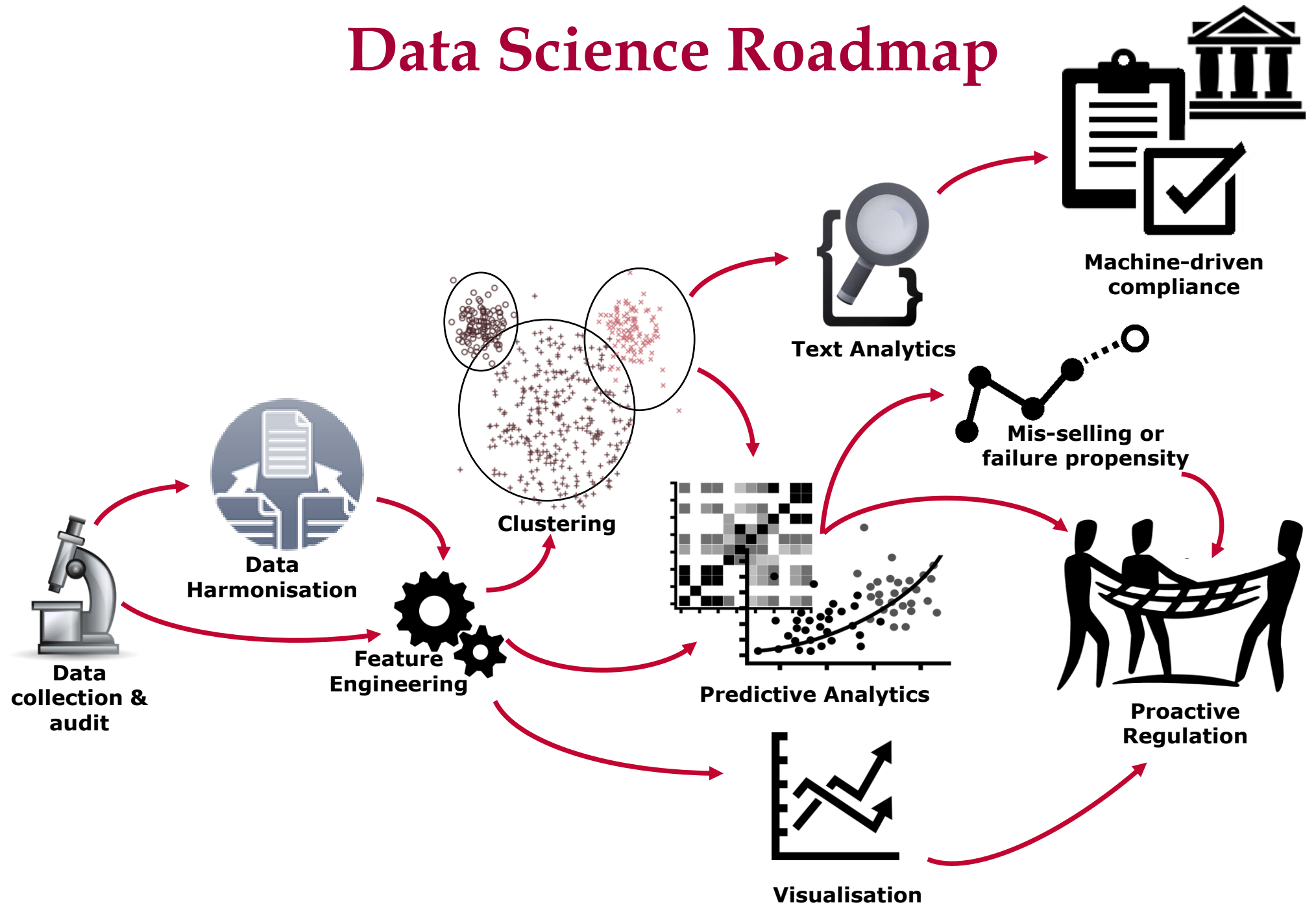
## *Retail:*

- 1. Impact of annual summaries, mobile banking and SMS alerts in **personal current accounts****
- 2. Field experiments on information disclosures in **savings** and **car and home insurance****

## *Wholesale:*

- 3. Impact of **high-frequency trading** on institutional investors**

# Data Science Roadmap



# Summary

- **DATA:** FCA collects rich transaction data + legal powers to gather more data
- **METHODS:** Undertaken rigorous, ground-breaking empirical research to inform policies. Starting to use range of data science methods
- **PEOPLE:** Empirical economists + data scientists
- **OPEN:** Open to new ideas for research + collaboration. Regularly work with world-leading academics + aim to publish in top journals
- **ACCESSIBLE:** Creating high-specification secure cloud environment facilitating off-site access
- **REAL-WORLD RELEVANT:** Research has to be immediately usable to inform policymakers

```
> print(`Thank you`)
```