UNIVERSITY OF EDINBURGH
Business School

CRC Credit Research Centre

# Highlights from Credit Research Centre

**Prof Galina Andreeva, Director**

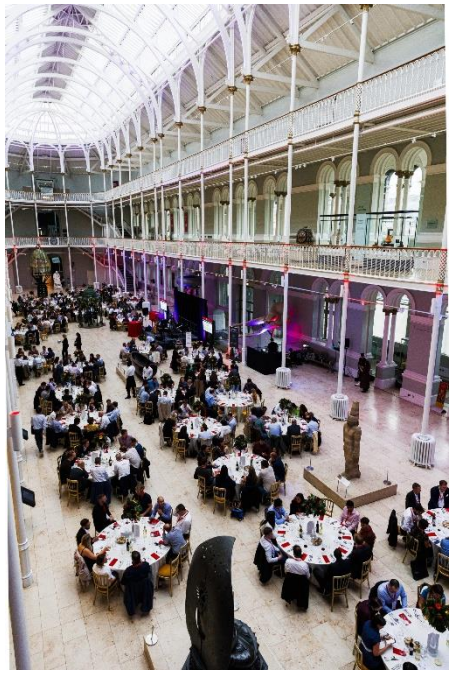
https://www.business-school.ed.ac.uk/staff/galina-andreeva

# Credit Scoring & Credit Control conference

The 19th conference will be in Edinburgh 27-29 August 2025.

Last conference in 2023:

- 400 delegates (20% academics, 80% practitioners),
- 34 countries: US to Australia and Norway to Vietnam,
- 9 commercial sponsors,
- 111 competitively selected papers in 5 streams,
- Special Issue of 'Annals of Operational Research'.



Registration by 12 August 2025

CRC Credit Research Centre

# Novel statistical learning approaches for open banking-type data

This thesis develops novel statistical learning methods motivated by open banking-type data as specified in European Commission (2018) legislation. Open banking legislation provides a framework under which, given the customer permission,financial institutions share certain information about customers, such as data on accounts and transactions, among themselves, and in some cases, with third parties. The scale and nature of this newly accessible data poses challenges for private, non-profit and government institutions. This thesis introduces new statistical methods that address some of the identified challenges in the analysis and use of open banking-type data. First, we propose an approach for comparing a specific reference subject (e.g. a customer of a financial institution or a user of a financial application) to a comparison group, defined via covariate truncation. The time-varying comparative mean value is developed here as a statistical method that learns about the dynamics governing how the response of a reference subject differs from that of a comparison group. We take advantage of local polynomial regression to provide a smooth estimator for the time-varying comparative mean value. By contrasting a specific customer against similar customers, our method offers interesting diagnostics that can be used by financial institutions to recommend personalised services, such as financial advice and products. Another contribution of this thesis is the development of a clustering/market segmentation method that takes advantage of the positive attributes of model-based clustering, while addressing some of its pitfalls. Standard model-based clustering is likely to result in the same number of clusters per margin. This seems to be a rather fictitious assumption for a wide range of situations to model. Model-based clustering is also computationally intensive for large datasets. The here developed Reign-and-Conquer clustering addresses both of these drawbacks by modelling margins separately, allowing for different number of clusters on each margin, whereupon the joint is reconstructed using an game theory inspired algorithm. Since the method specifies only marginal models, it is partially parallelisable and therefore computationally appealing for large datasets, such as open banking-type data. Practical application and overall good performance are demonstrated on both artificial and real data scenarios. This thesis concludes with a discussion of the here introduced methods, their potential extensions and opportunities for future research.

**PhD thesis by Andrej Svetlošák**

**https://era.ed.ac.uk/handle/1842/41562**

# Unleashing the potential of open banking: three use cases for risk assessment

Current account transactions offer an objective and up-to-date 360-degree view of an individual's financial status. This type of data can facilitate better access to financial products, and also answer a lot of important societal questions.

**Example applications/ use cases:**

PhD sponsored by EIT Digital and Atto, Ms Rui Goh https://era.ed.ac.uk/handle/1842/42242

**(1) Vulnerability-related volatility**

Volatility score describing the degree of uncertainty in cashflow projections
Identify target risk group (harmful volatility driven by vulnerability).
Working Papers | Credit Research Centre

**(2) Predict (impaired) cashflow**

Probability score to classify accounts into different risk levels
Identify the predictors of the financial distress.

**(3) Transient vulnerability**

Advanced latent state models to identify the hidden stages of financial distress, and transitions between them.

**(4) Exposure risk to suspicious activities**,

e.g. “shell” accounts
Model typical vs atypical behaviour, Outlier score to rank the risk.

“*Understanding consumer’s cashflow more effectively is of great importance to us as we seek to build products to assist both consumer and business. Open Banking has been shown to be the best way of understanding consumers based on their bank data.”*

James Varga, Atto Founder

# OB for Financial Vulnerability and Consumer Duty

Joint work with Professor Tina Harrison, sponsored by Salad Projects Ltd

> "Thousands of NHS workers have been left heavily reliant on several high-cost loans charging interest of up to 1,333% because they are being excluded from more affordable mainstream options, a new report claims. The study by researchers from the University of Edinburgh Business School examined the finances of almost 10,000 mainly lower-paid and younger NHS workers, and found that almost a third (30%) were using five or more loan providers, many of which were high-cost, such as payday and short-term loan firms."
>
> *The Guardian, 21.02.2021*

More information:
Harrison & Andreeva (2021) 'Financial Health of NHS Workers' Report 1
Harrison, Andreeva, Garzon-Rozo & Zhang (2022) 'Financial Resilience and Credit Landscape of Public Sector Workers' Report 2

Andreeva, Harrison, Rowllings & Zhang (2025) 'Developing Standardised Measures for Consumer Duty Compliance, Affordability and Competitiveness in Consumer Lending' Report 3

CRC Credit Research Centre

# THE DOUBLE-EDGED SWORD OF BIG DATA AND INFORMATION TECHNOLOGY FOR THE DISADVANTAGED: A CAUTIONARY TALE FROM OPEN BANKING

This paper examines the ethical and regulatory challenges arising from the advent of new technology and new data. We use an example of the 2nd Payment Services Directive (PSD2), known as Open Banking (OB) in the UK, however, the results serve as a warning to data expansion, in general, by showing how the new data can inadvertently correlate with protected personal characteristics. The Directive has dramatically changed the landscape of household finance by allowing access to real-time consumer bank data. While this facilitates financial innovation, it also raises significant concerns about privacy, fairness, and the potential for systemic biases. Utilizing a substantial dataset of 180 million bank transaction records, our study (1) demonstrates the power of OB data by providing the nuanced insights into financial vulnerability (FV) and (2) explores the above concerns, specifically focusing on how seemingly neutral OB transactions may conceal links to protected and sensitive personal characteristics. We develop FV indicators derived from OB data and use advanced machine learning techniques to explore FVI associations with financial behaviors and sensitive personal attributes. The findings emphasize the regulatory need for frameworks that adequately address the potential risks of modern data-driven financial services, ensuring that they are harnessed responsibly. Our findings underscore continuous specific needs to address nuanced challenges within financial services. This research contributes to the dialogue on ethical AI and financial regulation, providing insights that are crucial for policymakers in enhancing protections for all consumers, especially the most vulnerable.

**CRC Working Paper 01/2023 by Kim, Andreeva, Robatsos**

# Questions for discussion

- What are the most promising uses of OB data?
- What new definitions/measures do we need to support these uses?
- What challenges are you facing with traditional credit scoring?
- Where could alternative data enhance your decision models?
- How can we make Open Banking adoption smoother for lenders?