

October 2024



CODE ENCOUNTERS REPORT 4:

Credit risk decisions, mortgage lending and technological possibilities

**by Alison Wallace, David Beer, Roger Burrows,
Alexandra Ciocănel and James Cussens**



Acknowledgements

The research team would like to thank the Nuffield Foundation for funding the study. Warm thanks also go to the lenders, consultants, technology firms, brokers, and mortgage borrowers for their participation and support of this project, their help is much appreciated. We would like to thank Paula Higgins of the Homeowners Alliance, Robert Thickett of the Building Societies Association and Peter Williams for help with the recruitment of mortgage borrowers and mortgage lending experts. Catherine Dennison from the Nuffield Foundation provided valuable support and advice to see the project through. Lastly, the members of the Project Advisory Group were also helpful in recruiting for this study, providing additional support and insight along the way and valuable feedback on this report. Any errors are of course our own.

Nuffield Foundation

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare and Justice. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org



Contents

Acknowledgements	1
Contents	2
Executive Summary	4
Preamble	8
Chapter 1: Introduction	10
Research Methods	11
Report Structure	12
Chapter 2: Background	13
Introduction	13
Credit scoring	13
Bias and discrimination	14
Alternative data and automation	15
Human intervention	20
Consumer perspectives	20
Conclusion	20
Chapter 3: Technological possibilities	21
Introduction	21
Automation	21
Human involvement in decision making	25
Barriers to new technology	27
Conclusion	30
Chapter 4: Data possibilities	31
Introduction	31
Credit data	31
Banking data	34
Identity data	41
Other data sources	42
Conclusion	45
Chapter 5: Winners and losers of new technology in mortgage lending	46
Introduction	46
Open Banking data	46
Inclusionary practice	48
Bias and protected characteristics	50
Mortgage futures	53
Conclusion	53
Chapter 6: Discussion and conclusion	55
Recommendations	56
References	58
Appendix: Code Encounters project research methods	64

Executive Summary

The Code Encounters Nuffield Foundation funded project was undertaken by the University of York and the University of Bristol to examine the digital risk profiling tools that shape access to housing. These tools are increasingly adopting new sources of data and algorithmic processing and include tenant referencing tools in the private rented sector (PRS), affordability assessments in social housing and credit risk decisions in mortgage lending. The project ran from 2022 to 2024 and the findings are based on 122 in-depth interviews with people who produce, operate and are impacted by these various digital processes. This is the first UK study to gather multiple perspectives on the construction, operation and impact of digital tenant referencing tools in housing.

This report highlights key findings from 31 interviews with technology firms involved in aspects of credit risk decision making, mortgage lenders, brokers, consultants and borrowers in England.

Summary

- Credit bureaus and credit reference agencies (CRAs) collect data relevant to assess a person's creditworthiness but the data resources have changed over time, moving from community reputation, to income, status and default histories, but can now include new resources such as current account turnover, utility bills, stability of mobile phone numbers and email addresses, banking transaction data and digital footprint analysis.
- Large parts of the mortgage lending sector are already highly data driven and automated with mainstream lending using credit bureau information to verify customer supplied data and automated property valuations to make lending decisions.
- New technologies like Open Banking, where customers give permission for lenders to access banking account transactions, are evidently becoming increasingly common in other credit lending scenarios but to a lesser extent in mortgage lending,
- The interviews revealed that there are significant barriers to further automation and using new data resources that include large legacy IT systems, regulation, professional resistance and the importance and cost effectiveness of human oversight to detect fraud or assess lending in non-standard circumstances.
- The drive to automate lenders' back office operations was clear and may improve customer experience, increase speed, create efficiencies and the quality of decision-making. There were, however, mixed views on the strength of the business case for change and some lenders in niche mortgage markets made a virtue of human oversight and manual processes at the margins.

What is the background to the study?

The credit industry has adopted sophisticated data technologies for over a century as city populations grew and personal community reputation alone was insufficient to support business with strangers. This turned qualitative 'soft' insights into hard quantitative data. Drawing on this

expanded credit information, mortgage lending notably embraced data driven decisions and automated processes when credit scoring was adopted in the 1980s/1990s, aggregating insights from vast digital resources that reflected a person's creditworthiness.

Modern credit scoring technologies contributed to the growth of the subprime mortgage market by enabling lenders to offer higher interest rates to high-risk borrowers. This approach expanded homeownership opportunities but also led to increased defaults, particularly among low-income and minority groups.

Now the sector stands at a new turning point with increasing digital data sources and artificial intelligence (AI) changing what is possible. While credit scoring aimed to reduce bias in lending, historical practices such as redlining and ongoing issues with economic discrimination demonstrate that algorithmic models can still reinforce structural inequalities.

This report highlights findings from the Nuffield Foundation-funded Code Encounters study that provides an in-depth qualitative analysis of digital risk-profiling tools that govern access to the housing market. It is the first UK study to appraise these systems from multiple perspectives.

How are mortgage lenders using new data resources?

The interviews suggested that compared to other forms of credit, such as small business loans or short-term lending, the adoption of alternative sources of digital data to inform mortgage lending decisions is currently limited. Credit bureaus reportedly continue to expand the data resources available to lending institutions but lenders need the capacity to analyse and deploy the data.

The introduction of these new digital data sources, notably banking transaction data from Open Banking, aims to enhance credit assessments by providing a fuller picture of an individual's financial situation. Open Banking technology provides detailed and near-real-time insights into income and spending, potentially improving credit assessments, particularly for individuals with limited credit histories. Despite its potential, Open Banking faces challenges such as managing multiple accounts, accurate transaction categorization, and consumer trust, which may limit its widespread adoption. There are also alternative ways for lenders to gain similar insights from extracting data from bank statements to categorise and analyse, as well as using turnover data supplied by bureaus.

The credit industry continues to grapple with data quality issues, including errors and duplications, which can impact credit risk assessments. Efforts to improve data quality through de-duplication and advanced analytics are ongoing. Lenders increasingly use direct data sources like Companies House, Land Registry, and HMRC, while traditional spatial data and behavioural data from cookies are less commonly used.

While consumers generally accept the trade-off of detailed credit scoring and financial surveillance for access to credit, the interviews across the piece highlighted persistent concerns about privacy, data quality, and transparency, and the mandatory or discretionary nature of some data collection. These issues underscore the need for improved consumer awareness of how various data are used, regulatory oversight and consumer protection.

How is the mortgage market adopting further automation?

The mortgage industry faces a technological crossroads, struggling to fully embrace advancements due to legacy IT systems. An early adopter of automated decision making in the 1990s now these systems act as a break on the rapid uptake of more recent technologies. Despite the significant impact of credit scoring innovations, recent efforts focus on improving customer journeys and automating administrative tasks. AI is increasingly used to streamline processes and reduce costs, though complex or unique situations still require human intervention. The integration of Open Banking and other innovations, such as Open Finance and digital conveyancing, holds the potential to enhance credit risk assessments and the customer experience.

The current market is dominated by a few large lenders who leverage automation, while smaller lenders focus on niche markets and personalised customer support. Fintech firms are exploring alternative lending areas due to difficulties in engaging with mortgage lenders. Regulatory requirements and competitive pressures are shaping the pace of technological adoption, with delays in the home-buying process attributed to slow conveyancing and other external factors.

Who are the winners and losers in a digital mortgage market?

The global financial crisis underscored the importance of considering affordability in lending. Stricter regulations introduced in 2014 have made homeownership more challenging for lower-income individuals but have improved market sustainability and reduced the risk of financial shocks. Open Banking, which analyses transaction data beyond traditional credit history, has the potential to expand mortgage access for individuals with thin credit files, such as young people and migrants, but may also exclude those with poor financial management despite having sufficient income.

While automated systems and Open Banking technology can streamline lending processes, they may inadvertently exclude marginal borrowers or those with complex financial situations, leading to higher costs or more challenging terms. Addressing biases in algorithmic models and ensuring fairness in lending decisions are crucial to preventing the replication or amplification of existing inequalities. A balanced approach that combines technological efficiency with human discretion is essential for fair lending practices.

Conclusion

The credit assessment landscape is undergoing significant change due to advancements in technology and data integration. Automation and Open Banking are reshaping lending practices, presenting both opportunities and challenges. The industry's evolution towards more data-driven and automated processes must be managed with careful attention to fairness, transparency, and consumer protection to achieve positive outcomes.

However, the integration of granular data and automated profiling into mortgage lending introduces several ethical concerns. Privacy, fairness, and bias are major issues, as new classifications based on digital footprints can embed moral judgments into financial decision-making. This could disadvantage certain groups and exacerbate social inequalities, raising concerns about people's limited awareness of the need for people to actively manage

their digital data profiles'. These issues highlight the need for careful consideration of the implications of using detailed behavioural rather than just payment history data for credit assessments.

Recommendations

Across the Code Encounters project, we identified universal themes that need to be addressed as well as sector specific that require attention, including the following that are relevant to Government, those responsible for financial education, risk profiling technology firms, trade bodies, lenders, landlords and agents.

- 1. To make visible how data and algorithms have been used in each decision** Provide greater transparency in the way data is gathered from and about borrowers and clarity about how these data will be used.
- 2. To establish agreed guidelines on the appropriate use of algorithms for stakeholders within the sector and tenures** Provide guidance to lenders on how best to deploy algorithms and new data resources.
- 3. To produce guidance on the use of data and algorithms for tenants and borrowers** Increase public awareness of how they must manage their digital profiles, including banking transaction data, much in the same way as the importance of managing credit scores has permeated financial education and public consciousness.
- 4. To retain human oversight in decision making** Not all people fit algorithmic models so human oversight should be maintained to ensure fairness.
- 5. To ensure the explainability of decision making** Organisations must be able to fully articulate how a decision was reached, including the data used, where algorithms were involved and the human oversight of the outcome.
- 6. To ensure the retention of flexibility and individually tailored decision-making** We would suggest having a system in place in which the inputs into algorithmic processing can also be adapted to enable flexibility and to ensure that both input and outcomes remain flexible and adaptable to the individual being assessed.

These recommendations are discussed in more detail in our *Overarching summary report 1*.

Findings from this study have already been published in peer-review journals and all reports, papers and briefings are available to download from the project webpages <https://www.york.ac.uk/chp/housing-markets/code-encounters/>.

Preamble

This is the fourth and final report in a four-part series that explores the development, operation, and impact of digital risk profiling tools in England's mortgage and rental markets. This report - *Code Encounters Report 3: Data and automation in pre-tenancy affordability checks in social housing* - specifically focuses on social housing in England, offering insights into the unique challenges and implications of risk profiling within this tenure. The details of all four reports are as follows:

- *Code Encounters Report 1: Housing and algorithmic risk profiling in England - Overarching summary report* (2024) by David Beer, Alison Wallace, Roger Burrows, Alexandra Ciocănel and James Cussens. Centre for Housing Policy: University of York.
- *Code Encounters Report 2: Digital tenant risk profiling in England's private rented sector* (2024) By Alison Wallace, David Beer, Roger Burrows, Alexandra Ciocănel and James Cussens. Centre for Housing Policy: University of York.
- *Code Encounters Report 3: Data and automation in pre-tenancy affordability checks in social housing* by Alison Wallace, David Beer, Roger Burrows, Alexandra Ciocănel and James Cussens. Centre for Housing Policy: University of York.
- *Code Encounters Report 4: Credit risk decisions, mortgage lending and technological possibilities* (2024) By Alison Wallace, Alexandra Ciocănel, David Beer, Roger Burrows and James Cussens. Centre for Housing Policy: University of York.

The reports are based on a study aimed at:

- **Understanding tool production:** The study examines how credit rating agencies (CRAs), lenders and data analytics firms build credit scoring and tenant screening tools. It investigates the data used, how representative it is, the criteria for creating risk profiles, and how this information is communicated to key stakeholders.
- **Exploring deployment motivations:** The report assesses why housing professionals adopt these tools, focusing on policy, market, and regulatory influences and the risks they aim to manage across different sectors.
- **Examining practical deployment:** The study looks at how these tools are integrated into everyday housing practices, how they affect professional judgement, and how compliance with regulations like data protection and consumer standards is maintained.
- **Investigating perceptions and awareness:** It explores how well professionals and consumers understand these tools, their awareness of the data involved, and the impact on housing access.
- **Assessing market impact:** The study considers how professionals and consumers respond to the use of these tools, including potential exclusion or gaming of the system, and the balance between privacy and efficiency.
- **Considering policy implications:** It reflects on the regulatory and ethical issues raised by these tools and their broader role in shaping housing markets.

All four reports along with a literature review, briefings and a series of articles published in peer-review journals are available to download from the project webpages <https://www.york.ac.uk/chp/housing-markets/code-encounters/>.

Chapter 1: Introduction

The digital ‘challenger’ bank Monzo provides customers with a yearly summary of their spending called ‘Monzo Unwrapped,’ similar to Spotify’s ‘Spotify Unwrapped,’ which gives a rundown of your music listening habits, including your most played songs and favourite genres (Raffa, 2024). In December 2023, users of Twitter, now known as X, humorously expressed shock at the insights revealed by Monzo Unwrapped, which detailed their spending habits, such as frequent takeaways or visits to local pubs. Many were amused or embarrassed by their spending behaviour being highlighted in this way, with some questioning why Monzo would present them with such candid reflections. However, what if this detailed analysis of spending habits went beyond entertainment and influenced something as significant as access to housing? Our banking transaction data can reveal intimate details about our lives, including our social networks, subscriptions and financial behaviours – even identifying *OnlyFans* subscriptions. Open Banking technology allows such insights (Ciocănel *et al.*, 2024), with customer consent, to be shared with financial institutions and companies, including mortgage lenders. While lenders might be less concerned with your dining-out expenses, they could be very interested in your overall income versus expenditure balance or identifying ‘red flag’ spending, such as problematic gambling habits. Lenders are exploring how these new data sources and advanced technologies like artificial intelligence (AI) and machine learning (ML) can be used to assess creditworthiness or determine an individual’s risk profile.

This example shows how new technologies, like Open Banking, are changing the way financial services ‘see’ us (Fourcade and Healy, 2017). Open Banking allows financial institutions to gain deeper insights into our lives through our financial data, influencing how they judge our creditworthiness. According to Fourcade and Healy (2024), in their recent book *The Ordinal Society*, digital data categorisation is reshaping culture, economic competition and social hierarchies in profound ways. They describe this as a ‘classification situation’ (Fourcade and Healy, 2013), where automated systems sort people into hierarchies based on data, embedding moral judgments into these processes and often disadvantaging certain groups. This has led to more detailed classifications of people based on their digital presence, what Fourcade and Healy call ‘*eigencapital*’ - a form of digital social capital derived from our online activities. People are thus increasingly pressured to manage their digital profiles to maintain a positive image (Langley, 2014), effectively becoming ‘entrepreneurs of themselves.’ Those who struggle to navigate such systems, whom Fourcade and Healy (2024: 131) term the ‘lumpenscoretariat,’ often end up receiving lower-quality services, highlighting how these classifications can deepen social divides. These tools enable markets to understand individual behaviours in a more nuanced manner than was previously possible, rather than relying on broader statistical categories. As a result, managing one’s ‘*eigencapital*’ becomes essential, with algorithmic risk profiling shaping access to services, which raises ethical concerns around privacy, fairness and bias.

This report describes the findings from our Nuffield Foundation-funded *Code Encounters* study, which looks at how data and algorithms are influencing access to housing, with a focus on mortgage markets and the digital risk profiling tools lenders use to assess creditworthiness. Of all housing tenures (see Reports 2 and 3 for discussions on digital risk profiling in private and social renting), mortgage markets, upon which most owner-occupation depends, have the most experience using extensive data and technology, having relied on digitally-supported credit scoring for a long time. Now, they are at another technological juncture with the rise of AI, ML and access to ever more data sources.

As cities grew during the Industrial Revolution, credit bureaus emerged as repositories of information about individuals' creditworthiness, initially relying on subjective, qualitative data from local traders to assess a person's character, capital, and capacity to repay (Lauer, 2017). Over time, these repositories evolved into more (supposedly) objective, quantitative systems, recording details about debts, wealth and other attributes. Early credit bureaus were quick to adopt new technologies, moving from handwritten notes and ledgers to typed records, and then using telephony to share information with lenders. The adoption of computer technology in the 1960s marked a significant transformation.

In the US, an employee of the Fair Isaac Company (FICO), the instigators of modern credit scoring, noted that credit risk assessment was shifting from an art form to a more scientific process, moving away from qualitative judgments to quantitative, data-driven evaluations (Lauer, 2017). This led to public concern and House debates about the development of data surveillance, which saw government initiatives blocked despite private companies using new large mainframe computers to assemble large databases containing records of most American adults. The integration of computing technology expanded the insights lenders could draw from credit bureau data, leading to the birth of modern credit scoring. This system of evaluating an individual's financial capacity to repay long-term loans, such as mortgages, has a long history of using multiple data points and automated processes to profile people. While technologically enabled credit risk assessments were not necessarily seen as superior to human evaluations, they allowed for a greater volume of applications to be processed (Lauer, 2017). For most people seeking loans, the shift from personal, face-to-face assessments to statistical analysis of a person's 'data double' is now complete (Lyon, 2002).

Today, we are at another pivotal point where digital innovations are expanding the possibilities for assessing credit risk. New data sources, including detailed banking data from current account spending and online behaviour tracked through cookies, combined with AI and ML, are significantly enhancing the scope, granularity and insights available for credit risk decisions.

Our digital mortgage futures are still speculative, but this study explored the evolving landscape of mortgage markets in the context of new digital and algorithmic technologies. To understand these changes, the research looked beyond the code and technology to the entire 'socio-technical assemblage' (Kitchin, 2017) and the broader 'regime of recognition' (Amoore, 2020). The study examined the ecosystem of people, institutions, and contexts surrounding digital platforms and tools, exploring encounters with new data and automated systems from multiple perspectives. This included insights from firms designing software for credit risk decision-making, observations on how these new systems operate, and the experiences of mortgage borrowers affected by digital risk profiling.

Research Methods

This report is based on an analysis of 31 qualitative in-depth interviews with national stakeholders (n=3), technology firms (CR) (n=7), lenders (n=4), brokers (n=3), consultants (n=2) and mortgage borrowers (MB) (n=12) – a subset of the 122 interviews conducted as part of the wider study (across all tenures). It is important to note that there is much blurring between these categories, some lenders designed their own credit risk decision engines so were labelled as credit firms, offering insight into their technology, while some firms only designed systems for external clients. Some credit firms were also platform brokers, so again straddled the technology/lender divide. Regardless of their current role, many professional participants had been employed in various roles across the mortgage market and had decades of high-level

experience. These professional interviews were solicited with direct approaches and snowballing techniques. Mortgage borrowers were recruited with the assistance of the *Homeowners Alliance*, who circulated research invitations to members on the project's behalf. Borrowers were offered £20 Amazon vouchers as an incentive and thank you for supporting the research. Interviews ranged from 30 to 90 minutes and were conducted over Zoom or the telephone, and the resulting audio files were transcribed. The interviews were undertaken between 2022 and 2023, but mostly in the second half of 2023. The pace of technological change has been rapid with the hugely impactful *ChatGPT* emerging during the study period, recasting possibilities. The evidence presented here is, therefore, a point-in-time assessment of a fast-moving object.

A key limitation of the research is (despite our best efforts) the minimal engagement of large mainstream lenders, as those lenders who did participate represented a small proportion of mortgage lending. Asking about credit risk decisions and lenders' plans for using new data automation is considered highly commercially sensitive and although non-disclosure agreements were offered few larger lenders opted in. Engaging senior consultants and national stakeholders with decades of industry experience, often in big-name institutions, provided some coverage of the issues faced, but identifying how lenders and particularly their underwriting teams engage directly with old and new technology would have been beneficial. Nonetheless, the research provides an overview of key issues in the mortgage market at this technological juncture.

Further details are available in the appendix.

Report Structure

The report goes on to examine the current evidence on technology and mortgage lending in Chapter 2. Chapters 3 and 4 explore participants' perspectives on the potential benefits and limitations of new technologies and data options available to mortgage lenders. If the mortgage market embraces broader data sources and increased automation, this could reshape who gets approved or denied during mortgage risk assessments, creating new groups of winners and losers – this is an issue discussed in Chapter 5. Finally, Chapter 6 provides some brief concluding thoughts on the findings.

Chapter 2: Background

Introduction

The development of computer technology in the 1960s marked a pivotal transformation in the financial sector, particularly in the realm of credit assessment. As lenders and credit bureaus began to harness the power of data, the traditional, personal methods of evaluating creditworthiness gave way to more systematic, data-driven approaches. This shift not only redefined the criteria for lending but also reshaped the broader landscape of financial services. Credit scoring emerged as a radically innovative tool, moving beyond mere risk evaluation to influence various domains including mortgages, utilities and insurance. The evolution from localised credit assessments to expansive national data centres, and eventually to sophisticated algorithms and alternative data sources, highlights the significant impact of technology on financial practices. This chapter explores what we know already about the history, development, and implications of credit scoring, examining how it has altered risk management, market dynamics and the pursuit of fairness in lending.

Credit scoring

Starting in the 1960s, lenders and credit bureaus began to harness computer technology to manage their extensive databases more efficiently. This technological shift reduced the personal interactions traditionally involved in mortgage assessments. Instead of evaluating creditworthiness based on personal character, lenders started to focus on financial probabilities (Lauer, 2017). Local and regional credit bureaus were consolidated into large national data centres. This consolidation shifted the focus from assessing an individual's sense of moral obligations and responsibilities to evaluating financial metrics and data points. During this period, US lenders developed their own statistical models for assessing credit, but by the 1980s, the FICO score became the industry standard. This system used scorecards to integrate mortgage applicant information with credit history data from the credit bureaus (Lauer, 2017). These statistical models, initially used for credit scoring, were later applied to various other sectors, including utilities, telecommunications, insurance and even employment (Rona-Tas, 2017). Credit scoring facilitated the expansion of marketing opportunities by allowing businesses to segment populations for targeted financial offers. Consequently, the lending relationship evolved from a personal evaluation of creditworthiness to a continuous process of risk assessment. The significance of credit scoring was cemented when Fannie Mae and Freddie Mac, US government-sponsored mortgage enterprises that purchase loan portfolios and provide securitised funding to the market, endorsed its use to lenders.

The introduction of algorithmic modelling and data analysis has significantly affected the sectors in which they are used. For instance, in the insurance industry, new data resources have disrupted traditional risk pooling methods because premiums are now more accurately aligned with individual risk levels (Cevolini and Esposito, 2020). In the mortgage market, the rise of credit scoring, driven by advancements in technology and the use of ever more extensive data reserves, shifted the focus from merely minimising risk, to managing risk and maximising profitability (Lauer, 2017; Fourcade and Healy, 2013; Adkins, 2018). Under this system, high-risk borrowers are no longer simply denied loans. Instead, they are charged higher interest rates, which led to the emergence of the subprime mortgage market.

Credit scoring and the subprime market played a role in helping individuals with thin, poor or damaged credit histories rehabilitate their credit and contributed to the expansion of homeownership during the 1990s (Munro *et al.*, 2005; Keasey and Veronesi, 2012). However, this system also led to increased consumer defaults (Wyly *et al.*, 2009). Before mortgage market regulations were tightened following the global financial crisis, the FCA identified that individuals affected by higher interest rates and default risks were often low-income, older, working-class mortgage borrowers (FCA, 2012). Many of these borrowers are now stuck on unfavourable mortgage deals, unable to re-mortgage due to their original lenders failing and the loans being sold off. For some of these borrowers, especially those with arrears from elevated mortgage rates and subsequent increases, low credit scores have created a feedback loop that traps them in difficult 'classification situations' (Fourcade and Healy, 2013). This highlights how credit scoring systems can reinforce stratification in access to debt and housing (Sparkes, 2024).

Although there are fewer specific studies on UK credit scoring, lenders widely adopted the technology because it helped them segment the market and develop risk pricing. It became a crucial part of underwriting processes (Wainwright, 2011). The technology was also seen as a cost-effective way to handle large volumes of information. While it led to some de-skilling of the workforce, it provided a timely response to social and economic changes and served as an inter-organisational tool for sharing knowledge (Leyshon and Thrift, 1999).

Bias and discrimination

Debates about algorithmic decision-making often focus on concerns about embedding societal biases in automated models. However, issues related to technology, data, automation and mortgage markets have existed long before the rise of 'big data,' AI and ML. In the USA, there is extensive literature on how mortgage assessment technologies have had what is termed 'disparate impacts,' meaning some borrowers are more likely to receive financing than others. Historically, practises like 'redlining' - using cartographical technologies such as the Home Owners' Loan Corporation (HOLC) maps - excluded mortgage lending to applicants from low-income and predominantly Black neighbourhoods. This practice continues to affect urban patterns today (Brick Underground, 2015). Additionally, local and personal credit risk assessments often introduced discretion and bias, leading to regular refusals of loans or predatory lending practices for divorced or single women, people of colour and other minorities. These practices imposed harsher terms compared to mainstream mortgage markets (Lauer, 2017).

Credit scoring emerged as a response to such subjective biases, offering what was perceived as an objective assessment based on factual data (Lauer, 2017). It was seen as a technological solution to subjective discrimination. The adoption of credit scoring was accelerated by the US Equal Credit Opportunity Act of 1974, which aimed to protect against discrimination lawsuits against lenders. Nevertheless, discrimination based on economic calculations, racism and sexism persisted as issues. Even with the Community Reinvestment Act of 1977, which required US lenders to invest in communities where they take deposits to address spatial racial discrimination, accusations of racial discrimination against lenders continued. Since 2021, the Federal Justice Department has filed 11 lawsuits against lenders for disadvantaging people of colour through redlining practices (OPA, 2024).

There are concerns that credit scoring models just reinforce structural inequalities (Fourcade and Healy, 2013; Schmeckpeper *et al.*, 2021; Langley, 2014; 2009; Marron 2009) as it is not just about who does or does not get credit lines but what products are sold and the terms on which

they are offered. There have been multiple concerns about how credit scoring enabled the development of pricing to risk (Poon, 2008) and the disparate impact of the subprime mortgage market crisis on black and minority ethnic communities in the USA, who were steered towards these less attractive products even when eligible for better prime products (Reid *et al.*, 2017).

In the UK, the Equal Opportunities Commission endorsed credit scoring as a more objective tool for assessing creditworthiness (Leyshon and Thrift, 1999: 448). However, there is limited evidence regarding current patterns of lending and inequalities related to protected characteristics. UK banks tend to provide fewer loans in poorer areas, but without data on demand-side applications, it remains unclear whether loans are awarded unequally or if certain groups are disproportionately affected (Rae, 2015). Credit scoring was therefore seen as less about fair access to credit and more about closing information asymmetries between lender and borrower (Leyshon and Thrift 1999).

In the current period, bias in AI and ML modelling can arise from incomplete training data or the groupthink bias of developers who may come from a narrow segment of the population (Umoja Noble, 2018). The evidence is, however, mixed on the impact of these technologies on fair mortgage lending. Bono *et al.* (2021) provide evidence that switching to ML models improves the accuracy of identifying credit risk and does not necessarily amplify or reduce bias towards certain protected characteristics compared to traditional modelling approaches. However, Fuster *et al.* (2018) argue that using more sophisticated ML techniques for credit distribution alone can lead to more accurate assessments of credit risk and increase mortgage provision in the USA. Yet, this can also disadvantage Black and Hispanic borrowers compared to the white majority. Additionally, a study on fintech lending data found that the greater detail provided by Open Banking data (discussed below) allows for proxies for sensitive or prohibited characteristics, which can lead to indirect discrimination (Kim *et al.*, 2023). Therefore, the principle of fairness by unawareness - whereby ignoring protected characteristics is assumed to eliminate bias - was therefore considered untenable in this new era.

Alternative data and automation

Credit risk models are normally proprietary and often use data that is not transparent and is kept confidential to prevent manipulation of the system. For example, the FICO algorithm was initially opaque, but Arya *et al.* (2013) managed to ascertain that it included variables such as payment histories, credit applications, negative events like defaults and bankruptcies and income. Deville (2020) identifies public sources used in credit assessments, such as the electoral roll and adverse credit histories, which include county court judgments (CCJs) and insolvency and bankruptcy data. Lenders have different risk appetites, which they express by varying the weight they assign to these data points. Although there have been calls for greater transparency, scoring algorithms remain largely obscure and can be arbitrary (Citron and Pasquale, 2014). Pasquale (2015:215) describes these systems as 'black-boxed – unknowable, but with a significant influence over our lives, with data brokers 'creating new currencies for opportunity and attention.'

Credit bureau or (as more routinely known within the UK) credit referencing agency (CRA) (we will use the terms interchangeably) data is an important resource, as is internal banking data if people are applying for a mortgage from the same institution where they hold their current account. However, some people do not have a long and comprehensive history of using financial services and therefore have limited credit histories or 'thin files', rendering some customers, typically young people, those on low incomes or migrants, invisible (Experian, 2018).

As a result, lenders may struggle to assess these individuals' credit risk, missing out on potential sales opportunities and falling short of the industry's commitment to 'financial inclusion' (Hohnen *et al.*, 2021).

To address these challenges, the industry is increasingly incorporating a broader range of data sources, facilitated by digital technologies that make administrative and other data more accessible and tradable. For example, in the US FICO has begun using alternative data sources, such as mobile phone payments, public records and property data (Henry and Morris, 2018). In the UK, additional data sources now include behavioural scores related to credit usage patterns, current account turnover, internet and device data, as well as rent, utility and phone payments (Deville, 2020). Experian (2018) suggested using data from UCAS, the UK's university admissions service, to provide lenders with insights into a young person's risk profile, including details about the institution attended, A-level results and family employment history. Hohnen *et al.* (2021:38) emphasise that 'all data is credit data', suggesting that any type of data can potentially contribute to credit assessments. Further pushing the boundaries of alternative data, Berg *et al.* (2020) developed a credit scoring system for unsecured lending based on digital traces left by consumers when purchasing goods online. This system was shown to be as effective as traditional financial services data in predicting default, and even more powerful when combined with it. These digital traces include cookie data that reveal whether a customer arrived at a website via a price comparison site or an advert, if they made spelling errors, the type of mobile device they used (from the latest and most expensive to the oldest and cheapest models), and the timing of their purchases - all inferring a more or less savvy consumer. *Seon*, a global provider of alternative credit data, suggests that models based on digital footprints can enhance financial inclusion, improve accuracy and allow for more personalised credit offerings. Another global provider *Credolab* suggests that such approaches can reduce defaults and increase approval rates.

These alternative data resources are primarily used by new fintech start-ups rather than mainstream lenders, highlighting a divide between traditional and emerging financial services. Hurley and Adebayo (2016:149) caution that credit scoring systems, which once appeared objective, might shift towards reliance on inferences and 'creditworthiness by association'. In these models, various networks and affiliations are used as proxies for trustworthiness, which can benefit some borrowers. However, they also point out that, at least in the US, existing laws (at that time) were not sufficient to prevent potential discrimination against others when these alternative scoring methods were employed.

Deville (2020:9) provides a useful overview of the expanded data sources now used in credit risk assessments. This is summarised in Table 1. These sources include payment histories from a broader array of services, many of which are not directly related to traditional financial services. Additionally, assessments now consider behavioural patterns such as spending habits, spelling errors, shopping behaviours and indicators of social status, including the brand and version of devices used for making applications. The landscape of credit risk assessment is, however, evolving rapidly, integrating a diverse array of data sources beyond traditional financial information. In addition to the data summarised by Deville (2020), several recent innovations have emerged that further expand the scope of credit assessments.

One significant advancement is the expanded access to transaction data facilitated by Open Banking. This data is utilised by tools such as *FusionScore* and the *Financial Health Index* from *Account Score* and *Equifax*. These tools analyse behaviours observed in a customer's debit and credit transactions, shedding light on their attitudes toward spending and saving. This type of data can also act as an early warning system for lifestyle and financial status changes, providing

lenders with a more comprehensive and dynamic understanding of an individual's financial health.

Table 1: New data sources being used in UK credit risk assessment (adapted from Deville (2020:9))

Data Type	Description
Behaviour scores	Predictive scores derived from a borrower's credit usage patterns with a creditor. Potentially combines data relating to financial behaviour – such as repayment history, or changes in the credit limit to credit balance ratio - with non-financial information - changes in address or marital status, for instance. Particularly used in credit card lending.
Current account credit turnover (CATO)	Reported by retail banks, drawing on income and expenditure information from users' current accounts. Uses algorithmic modelling to estimate borrower incomes. This is a challenging task, according to a number of interviewees, giving the complexities of many individuals' financial arrangements. Access to data varied according to the POR, but even those lenders not reporting CATO data can employ services offered by CRAs to undertake income verification.
Internet and device data	Digital information released by a user accessing an app or webpage. Includes IP address, which is used to estimate a user's location, and information about the type of device (e.g. make, model, operating system) being used. Used by at least one credit reference agency in their fraud prevention work.
Omni channel	Using data derived from communications with borrowers across various 'channels' – for example phone, email, SMS. Can allow a more dynamic assessment of risk, in which risk is determined by patterns in how a user responds to a prompt via a particular channel. Reported to be being used by some creditors in the UK and more widely overseas.
Phone and broadband	Data concerning payment history with communications providers. A new addition to CRA data.
Social housing rent payments	Data reported to CRAs by social housing providers regarding tenant payment histories and listed in credit reports. A new addition to certain CRA data.
Private rented sector payments	Data reported to a CRA by either an individual tenant or a property management organisation. Currently restricted to one CRA, via Experian's recently launched 'The Rental Exchange.' This allows individual private tenants and large property management organisations (letting more than 500 properties) to have private rental payment data recorded on an individual's credit file. For most private renters it remains an 'opt-in' possibility.
Utilities	Data concerning payment history with utilities providers. A new addition to CRA data.
Vehicle finance	Data concerning payment history with vehicle finance providers. A new addition to CRA data.

Despite its potential, the use of social media activity in credit assessments has seen low adoption in the financial sector and faces high levels of scepticism. A few years ago, a CRA and a major bank explored the integration of social media data into their credit models. However, they found that it did not sufficiently enhance model accuracy to justify continued use (CDEI, 2020). Similarly, in 2016, *Score Assured* launched *Tenant Assured*, which requested access to tenants' social media accounts for application assessments. This initiative was later withdrawn due to privacy concerns and the limited effectiveness of social media data in improving credit assessments (Ferrerri and Sanyal, 2021).

In certain markets, particularly emerging ones, fintech companies are experimenting with mobile phone usage data in credit scoring. These models consider variables such as call duration, the timing of calls, frequently contacted numbers, call initiators, and even the grammar and punctuation used in text messages (Berg et al., 2019). These data points are intended to provide insights into a consumer's habits and reliability, offering alternative perspectives on creditworthiness.

Credit assessments also incorporate data related to payment behaviours on subprime loans, such as payday loans. This information provides additional context about a consumer's financial habits and risk levels, potentially revealing patterns that are not captured by traditional credit data (Foohey and Greene, 2021).

While transaction data has gained widespread acceptance and is valued for its insights, other innovative data sources - like social media activity and mobile phone usage - face mixed receptions and are predominantly applied in niche or emerging markets. The ongoing challenge lies in leveraging these alternative data sources effectively while addressing concerns about privacy, accuracy and their relevance in various financial contexts.

What is clear, however, is that Open Banking technology represents the likely most significant new data resources available to lenders, providing detailed current account transaction data through third-party agencies. This technology was introduced in Europe under the European Commission's revised Payment Services Directive (PSD2) (Kassab and Laplante, 2022). Open Banking allows firms to 'automate and simplify the interpretation of bank-statement data,' which helps assess a person's affordability and propensity to pay (Experian, n.d.). Additionally, Open Banking supports open payments, reducing the reliance on third-party card services that typically facilitate payments between different entities. However, the primary focus here is on its use for analysing bank account transactional data.

Open Banking is being utilised across various scenarios, such as providing affordability assessments, automatically saving spare cash (e.g. Plum), setting spending limits on gambling apps (e.g. True Layer) and applying for credit lines (e.g. Salad Money). This technology uses APIs (application programming interfaces) to connect with other credit risk decision-making platforms, granting access to an individual's current account data for up to 12 months prior (and three years for business accounts), and for a period of 90 days following consent. Typically, firms work with a Financial Conduct Authority (FCA) licensed third-party broker, which could be a CRA or a smaller intermediary, responsible for accessing these data in compliance with regulatory standards.

These brokers or data aggregators enhance the value of this data by providing various analyses alongside individual transaction data. This includes details of each income and expenditure item, including the timing of transactions, as well as ML applications that categorise transactions into various spending categories such as utilities, food, entertainment and gambling and

distinguish between fixed and discretionary spending. It also categorises income types such as employment, benefits, or money transfers from other accounts. By aggregating data across multiple accounts, brokers can create a detailed portrait of actual financial behaviour and provide insights into a person's lifestyle.

The timing of transactions is also recorded, allowing analysis of spending velocity - how quickly individuals deplete their account balance after being paid - which helps infer positive or negative financial behaviours. A critical aspect of Open Banking is the requirement for account holder consent, which is fundamental as the concept of Open Banking was partly founded on the principle that individuals should have greater ownership and control over their financial data (Reynolds, 2021).

Mortgage market regulation requires lenders to undertake detailed affordability tests and Open Banking is seen by some to be an effective tool to support this activity. As in social housing, how lenders assessed expenditure varied considerably. An Equifax (2024) survey of 127 lenders in 2022 (see: <https://shorturl.at/Sp9cy>) found that 24% of organisations relied on a customer's self-declaration to assess expenditure; 19% relied on a CRA as their primary data source to assess expenditure, and 16% relied on Office for National Statistics (ONS) data to assess expenditure. The remaining 41% of lenders used a wider array of techniques in smaller proportions. Equifax can provide lenders with ONS and Minimum Income Standards data (an annual consensus-based estimation of expenses to meet a minimum acceptable living standard) but found ONS data was 25% higher than data revealed when using Open Banking data. Customer's recall of their expenditure is poor, according to their survey data, and ONS data is out of date when first used and may not reflect any rising costs of living. Equifax found that the addition of Open Banking data combined with the other data sources was 39% more predictive of customer defaults.

A further Salad Money (2023) report indicates that traditional credit scoring and information works well for mainstream markets but underserves those at the margins or those with thin credit files. In situations where people manage their finances well but cannot access more affordable or high street lending, Open Banking can provide an accurate and more inclusive financial appraisal that facilitates lending to people currently invisible to financial services.

Fintech companies have widely adopted algorithmic credit scoring, leveraging a broader range of data enabled by digitalisation and advanced server capacities. They are also utilising more sophisticated modelling techniques in lending, such as complex ML tools like random forests and artificial neural networks (Aggarwal, 2021). However, these advancements have not yet fully permeated the mortgage market, where logistic regression models remain prevalent, mainly due to regulatory compliance and the need for model explainability. Nonetheless, the literature indicates a growing trend towards ML applications in credit scoring models, along with diverse techniques to justify decision-making (Dastile *et al.*, 2020; Bucker *et al.*, 2022; Onay and Ozurk, 2018). In mortgage lending, ML is most commonly applied in areas like risk management and compliance, customer engagement, credit, securities sales and trading and general insurance. The primary perceived benefits of these applications include anti-money laundering, fraud detection, and overall efficiency improvements, which come with associated cost savings (BoE and FCA, 2019). However, the sector identifies key obstacles to innovation, including challenges with data availability, quality, and ethical usage; the need for sufficient explainability; a risk-averse culture; and difficulties in assessing consumer and public acceptance (CDEI, 2020).

Human intervention

Lauer (2017) highlights that automated decision-making was not seen as superior to the human underwriter, just more efficient, enabling expansion and competitive advantage so a larger volume of applicants could be processed in each period. Nonetheless, credit managers who were skilled at determining risk and underwriting felt threatened, as engineers had to convince 'Luddite' credit managers to adopt automated credit scoring (Marron, 2009). In this current period of technological innovation, financial service regulators found that UK firms had adopted AI in ways that continued rules-based models, augmenting current analytic techniques and leaving in place a role for human decision-making in most cases (BoE/FCA, 2019). Human insight, interpretation and decision-making remains.

Consumer perspectives

In the 1970s, concerns were raised about the *Bank of America* using credit information to segment the public and pre-qualify customers for credit cards. This issue led to hearings on Capitol Hill, where it was argued that linking payment history data with marketing practices had normalised the acceptance of routine invasions of privacy (Lauer, 2017: 267). Such technologies contributed to the normalisation of opaque algorithms that 'adjudicate everyday life', and the customer from that time would likely be astonished by the level of insight provided by modern Open Banking technologies.

Despite these concerns, consumers have largely accepted detailed credit scoring and financial surveillance as a trade-off for obtaining goods and loans, prioritising their immediate needs over privacy. This acceptance persists even though credit assessments make inferences about a person's honesty and responsibility, effectively providing a mark of shame that classifies individuals based on perceived extravagance or weakness (Lauer, 2017: 268). Rather than simply reflecting economic risk, credit scores often reify character and moral judgments about individuals as consumers, workers, or citizens. A recent study by the FCA (2023) on the credit information market highlighted issues with data quality. It revealed that consumers are often unclear about how to correct errors and resolve disputes with Credit Reference Agencies (CRAs).

Conclusion

Lenders have a long history of using financial services data and publicly available information to make credit risk decisions, a practice facilitated by advances in computing technology. Recent innovations in computing have expanded the range of data used in credit assessments to include a variety of non-financial information. AI and ML now enable the processing of this data to produce faster and more detailed insights. Consumers may not always be aware of these advancements. There are concerns that new algorithmic models could inadvertently reinforce existing social inequalities. However, the evidence on whether these issues specifically impact credit scoring remains mixed.

Chapter 3: Technological possibilities

Introduction

A lending consultant we spoke with, identified three types of automation: process automation; credit scoring; and algorithm-based decision-making. This chapter focuses on the project findings that relate to automating processes and lending decisions, while the next chapter will cover credit scoring and the use of new data. The findings draw on the literature that highlights issues of new data, automation and human insights as well as impacts on different groups, themes explored over the following chapters.

In the mortgage market, using data and automation for credit risk decisions has a long history, unlike in social or private rented housing. The industry is now at another technological crossroads but appears hesitant to embrace change fully. Credit scoring, which emerged in the 1990s, significantly transformed the market and led to innovations like subprime lending and securitisation. However, lenders have struggled to update their existing IT systems, which is more challenging than starting anew, as seen with agile fintech companies in other lending areas. The outdated technology in traditional firms has slowed further digital advancements. While there are ongoing changes, the full potential of technological innovations in credit risk decisions, as promised by technology firms, has yet to be fully realised in mortgage lending, though progress is being made.

The chapter begins by exploring which specific processes or decisions software companies and lenders are currently targeting for automation. It then looks at situations where human involvement remains important in mortgage lending. Finally, it discusses the varying rates of technological adoption across the lending industry and the challenges mortgage lenders face in implementing new technologies.

Automation

Credit scoring has already automated much of mortgage lending, enabling lenders to advance loans at volume for mainstream borrowers and properties and that has been consistent for many years. As discussed in the following chapter, credit scoring and the related computing power enabled the fast processing of a range of credit bureau and CRA data and underpinned the expansion of the market. Beyond this widespread adoption of credit scoring and automation from the 1990s in the UK, the pace of change has slowed and now the industry lags behind other financial services.

'The top six [lenders] do 70% of the lending. They have completely industrialised the process of mortgage lending. Their products are not innovative, but they can process really quickly and easily. It's an industrialisation to cut costs and do volume.' (Lender 7)

'I would argue that all of those people are very biased towards telling you and everybody else about fantastic new technologies and improvements in credit scoring that has occurred over the course of the last few years. My contention is, I've been working pretty closely for 25 years. And there's not been a lot of change.' (Lender 4 consultant)

'Mortgages haven't changed that much. That's because of the governance and the process around it, and the length of time it takes to make a decision. If you look at, say anything other

than secure lending, so auto lending, some lending, any retail products at all really, that falls under financial services, have all massively transformed. [...] I might be able to get a decision in principle relatively quickly, but the actual flow of being able to make a decision and moving that through from from "I like this house, I want to buy it" to actually buying it is probably the same as it was, you know, 20 years ago, 25 years ago.' (CR7)

Credit scoring and related technology have allowed data to be processed quickly. However, administrative tasks - such as customers submitting applications, staff handling these applications, and entering data into the models - have not advanced as rapidly. Recently, mortgage lenders have concentrated more on enhancing the digital experience for customers, like offering online application forms, rather than fully digitising internal administrative processes. Currently, many lenders are focusing on speeding up these back-office tasks, automating processes to streamline the customer journey and aiming to boost efficiency, save staff time, and reduce costs.

'The banks have spent their investment dollars on making things slick for customers, not making things slick for bank employees. So as a bank employee trying to deal with this, there will be system support. But a lot of these systems are a bit rubbish, you know, day and so on. And it's not uncommon [to] actually sort of manually key between systems and all this kind of thing. [...] So there's automation of the process, which basically means to what extent is it almost a straight through process where the customer is interacting digitally online, or on the map, or something like that. And the machine is whirring away in the background. And all these things are happening. Without a human being touching it in sort of let's say in real time, right? And there has been revolution in that in, over the course of my career. Not to say that all processes are automated, but everybody is moving in that direction. You can see a very clear trend, and there are more advanced players that can get you to a sort of decision in principle. [...] There are still people, particularly on the sort of private banking [...] right is still largely manual. And there's every permutation combination in between. But importantly, that, that's about the process, not how the decision is made.' (Lender 4 consultant)

'I think the question is to what's changed massively, from then to today is, that is the speed of is the amount of data that's available. And this and the speed at which we can get access to that data and make it valuable in a real time decision. 20 years ago, there was a big push called big data and big data was the thing. 'Big data is coming. Be ready for big data.' There's gonna be more data than you could [need]. What you need to do is go out and find the data that you want. And now that big data isn't, big data isn't the problem. It's, we're in big data. The problem now is there's so much data that's readily available. The skill is being able to access that data and get the value from it that you need to be able to make an accurate decision in less than a tenth of a second. Otherwise, the customer's going to go somewhere else.' (CR7)

'We work in what they call an agile technology... we need to have a culture where changes are constant. [...] We are not at the front of the game if there is a game, but we are very forward-thinking. We work in an agile technology, and that's how we carry on, with 3-week deliveries of small changes. [...] We've got a lot more to do, we've got departments which need a lot of focus, that need technology change.' (Lender 6)

Some lenders are automating their administrative processes by organising workflows between teams, automatically generating emails to request information or update applicants on the status of their application, sending relevant SMS messages, and even allowing applicants to click to accept loan offers formally. One company provides lenders with generative AI technology (similar to *ChatGPT*) that can instantly pull information from different parts of the management information system. This AI helps to fill in call handlers' screens, produce loan documents, attach appendices, and make formal loan offers immediately. Other lenders are exploring technology to save time and costs in reviewing documentary evidence needed for loan

applications, while a lender specialising in emergency short-term loans highlighted the advantages of automating loan decision-making processes.

'Yeah, so advantages, it's all about speed. So therefore, then, I'm removing the need for a manual underwriting. So, if I'm looking at bank statements that could be removing half day out of it instead of, could you just check this because we found something untoward within the bank statement. So that's a lot of what materialises. Obviously, that's twofold, because that can then really help with more business coming through for the same amount of people that I've got there. So, there's the efficiency gains.' (Lender 5)

'The automation is certainly something that a lot of the mortgage teams are kind of keen on. And it's potentially not from a credit decisioning perspective, but it's more from a workflow of how do they get and to a decision faster.' (Lender 3, consultant)

'We've now created our own AI machine learning model based on a machine ingesting all of the categorised data. Then, creating a score. We've created our own internal score for these consumers. Yes, basically, something like 82 per cent of the people that apply to us are decided now by a machine. [...] What's interesting is that you get, for most consumers, between 4000 and 6000 data points over a 2-year period [Open Banking data]. [...] What we do is we pour that into our machine learning/AI. You could argue whether it's AI or machine learning, I think it's a combination of the two. The system, then, looks at all of those points that differentiate a good payer from a bad payer and comes up with a score. We put that in last August and it has probably halved our loss rates which is quite staggering. We're now using that to auto-accept customers but also to auto-decline customers. Where our staff were looking at each and every example.' (CR4, short-term loan lender with own technology)

Automation at scale can simplify decision-making, but it struggles with handling marginal or unusual cases. Adding more stages and integrating extra data into the process can be costly, but it allows for customising data requests based on individual customer needs. For example, lenders can design the system to add extra steps for customers who are likely to fail affordability checks, prompting only these customers to consent to Open Banking. This approach avoids asking all applicants for additional information that might not be relevant to most lending decisions, thus creating a more personalised and efficient process.

'Companies strive to have straight-through processing. So, where they can remove manual intervention or manual referrals, they'll do as much as they can to do that. Now, in a housing market, the easiest way to try and straight through the process is to try and get all the information upfront, then have a computer say yes or no. If you're building in interim steps that try it one way and then actually, we don't have sufficient data, let's try this way, then the process you're building has multiple steps and is often at odds to strive for a straight-through process.' (CR6)

Beyond automating administrative tasks, the market now also offers credit risk decision-making software that gives lenders more control and flexibility to respond to external changes as needed. Traditionally, lenders would buy IT systems that were rigid and could only be adjusted by the technology provider, limiting the ability of internal staff to make changes. Now, companies are offering 'software as a service' (SaaS), 'low code' platforms that lenders can easily customise. These platforms allow lenders to adjust their own scorecards, models, and lending criteria directly, without needing to rely on technology providers to make changes, providing greater agility and responsiveness.

'So, the end game is for clients to be self-sufficient in a manner exactly as you've described. When clients first purchase [name of firm], we do hold their hand. We do train them. We may build it for them to start with, but the idea is that they can become self-sufficient, which does a few things. It allows them to not necessarily incur huge costs in having somebody else do the work for

them, but it also means that they can react to the speed of the market. So, if they need to make a change to how they built a decision flow... Maybe there's a new piece of data that's come that they want to use in that, they can change that flow and build it themselves, as opposed to having to go back to a consultancy firm to redo it and it takes months, which talks to the whole premise of [name of firm]. So, [name of firm], as you say, is a low-code risk decision platform. So, what we allow our clients to do is build their credit policy and risk decisioning [sic] flows on a [name of firm] studio.' (CR6)

'So quite often, a lot of our clients that move to us, they have hardcoded decision engines. So, they know it isn't their job to go and change that. And if that platform is more than a few years old, the person who probably built the platform has got another job somewhere. So, to unwind that and get someone else to come and maintain it is really challenging. So changing people's decision engines is actually a really hard thing to do for lots of different reasons that inquiry, lenders are cautious about changing it because unless you've got a good analytics team you don't necessarily know the consequences. Is that going to increase our decline rate? Is it going to increase acceptance rate? Is it going to increase our arrears rate or first payment default rate? Sure, the ADP [auto decision platform] is a self-service tool. So, the vast majority of what they need to change. It's, it's a real time tool because they can log in as a service, so it's not even software you traditionally install on servers they control or locally, so it's a service that we provide. Each of our clients has their own individual ADP instance. So, they control exactly what goes on there.' (CR1)

One emergency lender is using AI to automatically send responses to applicants who are declined a loan. This is possible because their AI model is designed to be explainable, meaning it can clearly outline the reasons behind the decision to decline the loan. This allows the lender to provide applicants with immediate, transparent feedback on why their application was not approved.

'ChatGPT...can write a very nice letter for you about, "I'm sorry you weren't able to get a loan because your gambling was too high. You should try and remediate yourself through GamCare/GAMSTOP." We are increasingly using... The largest amount of inbound emails is from our decline population. We're using AI to write back to those individuals based on their exact circumstances to tell them how they can turn that around. We know exactly why [loan declined in the model], yes. None of it is subjective.' (CR4)

The level of automation in mortgage lenders' administrative and decision-making processes varies depending on the market segment they operate in and their target customers. Automation is more straightforward in the mid-market, which deals with typical borrowers and properties. However, it's more challenging in markets involving high-value individuals or properties with unique features or applicants with complex financial situations. Automation is also difficult for lower-income households with unstable incomes. These scenarios often require more manual human intervention, which is discussed further in the next section and in Chapter 5.

'It depends on the nature of the lender as to what ... kind of the products that they're offering as to whether they're going for kind of more vanilla mortgages. So, it's kind of mainstream people that you're buying a 2-bed terrace. You've got a standard income, standard process. It's not a complicated thing. It's quite automated. Where you get some more niche lending...high value customers...that's where it becomes more mainly, underwritten and there's an advantage of requesting certain proofs of income, because you can't really get standardised, so difficult to automate it, because you could have people with random bonuses, income from multiple sources. So, a lot of it depends on the proposition of the client as to what they're kind of looking for.' (Lender 3 consultant)

The mortgage market sees the role of intermediaries or brokers as an area ripe for automation, given that brokers handle most mortgage applications. A consultant suggested that brokers mainly perform basic data collection and administrative tasks, which could be easily replaced by technology. Some companies have entered this space, offering online platforms for younger clients to complete fact-finding tasks that are then compared against comprehensive product databases. However, these technology-driven platforms still need backroom staff to support customers. Traditional, in-person mortgage brokers believe that effective fact-finding from clients requires human interaction, as certain issues - like a client being pregnant - might only surface in natural conversations and could be missed on application forms. Borrowers also value the personalised guidance or 'hand-holding' that human brokers provide, which a purely automated system cannot replicate in a meaningful way.

Human involvement in decision making

A common benefit of automation across all housing tenures examined in the broader project is its ability to free up time for human intervention. This allows staff to focus on areas where algorithms are less effective, such as conducting sense-checks, making nuanced decisions or providing detailed customer support, as well as decision making. Technology firms recognise that humans have an intuition or 'nose' for certain situations that are hard to replicate with code. As a result, automation is often designed not to replace human roles like underwriters, but to allow them to be used more strategically and effectively where their expertise is most needed.

'Algorithm-based decision-making...as much as everybody will tell you yes, the truth is, in most major providers, there's still a human being involved.' (Lender 4, consultant)

'And so, you've got in theory the underwriter, looking at the risk of the case rather than chasing bits of paper.' (Lender 6)

Manual underwriting remains crucial in the mortgage market, particularly for tasks like fraud detection where AI can only flag potential risks. For high-value transactions or complex financial situations, human underwriters are needed to navigate complexities that some fintech solutions struggle with. The approach to credit decisioning varies by lender and their products, but the cost of employing staff to review what automated models are doing is considered worthwhile when weighed against the value of the loans and potential losses in mortgage lending. Automated systems are valuable for highlighting issues that require human review, ensuring that many lenders maintain human oversight. For example, some lenders find it beneficial to have staff contact customers directly to verify details, detect issues like calls being routed overseas, or identify signs of problems such as intoxication. While automated models handle straightforward cases - like the 'vanilla' lending that makes up the majority of large mortgage lenders' portfolios - underwriting teams remain essential for managing non-standard cases and mitigating fraud risk.

'You could prove pretty quickly that credit model will rank all the credit worthiness better than the human being on average. The leap from there to "we should make the decision based on the credit score" is a bit of a naive one. So, what you're looking to do with human being in this process is to stop yourself from making stupid decisions that is somewhere between credit risk, and fraud. You know. Often there are non-standard things that the customers coming in with, etc., etc. So, most people will still have a human being apart, a little bit of time to it, and particularly with a mortgage, you know, if you're gonna be putting half a million out of the door. If a human being looking at something for half an hour, how much is a human being cost? You know, maybe 100,000 pounds a year or something, that human being, you know, if we're lending 500,000

pounds once every 5 years, if they spot something. It's a bit funny. They've kind of made that salary.' (Lender 4 consultant)

'We've got a lot of the grey areas already built into our technology, so we can offer people these options. Fundamentally, this is one of the issues you have with fintech lenders, which have tried to fully automate - there aren't very many of them, if any - is that they are very black and white. It's either yes or no. Fundamentally, all lenders, all the mainstream lenders are looking at manual-underwriting complex cases, and that will be, they normally have a number of senior underwriters working in the place, and they will assess the case based on, I think it's they've got a risk-appetite for it. So, you have a more shiny front end, but fundamentally, there is someone taking a judgement on it. These are relatively large amounts of money, so you can understand why they're doing it. They might then add a little premium if they need to, but yes, there's a lot of manual underwriting for the more complex or less standard cases that happens.' (CR3)

Participants acknowledged the value of manual underwriting but also noted that human intervention is not always beneficial. In some cases, human involvement could lead to inappropriate lending or less equitable decisions, as highlighted by past discriminatory practices reported by one technology firm. One lender noted that while the quality of human underwriting was comparable to that of automated systems, the machine was able to identify variables that human underwriters might miss. Additionally, the machine's reliability, efficiency, and cost-effectiveness often made it a more favourable option compared to manual processes.

'If you've got an experienced underwriter, they should be able to beat the scorecard or the strategy that's assigned to it, because that's, that's their job. The scorecard's been trained on a generic population, whereas an individual consumer's situation should be different. And you can kind of pick apart good and bad things.' (Lender 3, consultant)

'So quite often when we looked at loans that went bad with big losses, they would be big and complex loans that had usually been very, typically, they'd been declined by the credit team and then the broker, and then the broker would say, and I'd tell you, this is absolutely true. The broker would complain, saying, uh, well, this is a fantastic job, Mr. Smith. I've known him for years. And how dare you turn this down? And he would appeal to the chief executive of the lender. So, when he looked at the incidence of default on large loans that, you know, you lost a lot of money on, quite often it was because there's a referral up to the line, and the CEO did it because he wanted to preserve the relationship with the broker. But actually, the right decision was to have turned it down.' (Lender 7)

'When I talk to people that really want the human element like removing automation to a part and going back to mass... Human element is almost like a 'back in my day' type statement. And humans, you know, going in. See? Your bank manager sounds great, but it's hugely flawed. Right? Do you get on with them? Are you wearing the right suit? Are you the sort of person that is comfortable going to see your bank manager? Are you going to get [a] fair and equitable decision? Did they see you in the pub last Tuesday, getting drunk with you, mate. It's a bunch of stuff that just means it's not as fair and equitable as you think.' (CR7)

'Interestingly, the decisioning that's done by the machine and the decisioning that's done by staff is about on par. We're obviously looking to see that people can service an additional debt and our, largely female, staff are as good as the machine at decisioning... Our losses have been halved because I think we are acutely aware of things like Direct Debit performance now. It's how... That's something that the machine threw up that you hadn't been looking at so closely. Well, we had, but it was more able to consistently pick consumers that are... It's about 2000 applications a day and it's very consistent. It's very measured.' (CR4)

In 2022, 84% of mortgage lending came through market intermediaries, such as mortgage brokers, rather than direct applications to lenders (IMLA, 2022). National mortgage broker firms

that emerged decades ago, leveraging telephone technology to reach broader markets, remain significant today. Currently, digital fintech brokers have emerged, offering platforms that automate the inquiry, application, and product selection processes. These platforms cater to younger clients who prefer digital services. However, despite their automation, these platforms have found that substantial back-office staff are still needed to manage and streamline processes. Consequently, broker platforms now emphasise their ability to provide human customer support. Borrowers interviewed for this project valued human interactions for navigating what they found to be complex mortgage applications. Previous studies have highlighted the importance of these personal relationships, especially for first-time buyers (Wallace *et al.*, 2022).

'I quite like the human touch. I like the fact that I had somebody there to speak to if I was struggling and to talk me through the full process... I don't think I'd want to just go the full automation [...] I found it a bit daunting, I think, because I was doing it on my own. So, it was nice that I had somebody there. So, I don't think I'd want to lose that.' (MB4)

'There are some factors that can't be quantified. There's some factors that unless I do come explain for you by myself, you won't really gauge whatever I did attempt to tell you.' (MB2)

'Having a human to guide you through it was helpful. I felt like he was on my side because I think a lot... I've had quite a lot of anxiety around trying to get a mortgage for the last five years [...] I'm not sure, you know. I'm sure a lot of things can be replaced by internet and technology... but having a human holding your hand and guiding you through it has been helpful.' (MB1)

One borrower highlighted that the broker allowed document submission by email which she found convenient, whereas the conveyancer had a digital platform that she found did not work effectively and where she resorted to email to liaise with the solicitor. However, this does raise security issues regarding the security of key documents on open email platforms.

The mortgage market has become dominated by a few large lenders, with three major players accounting for 43% of all mortgage lending. Due to the high volume of lending, these large lenders rely heavily on automation. In contrast, some smaller lenders focus on niche markets or specific customer groups that the large lenders may not serve effectively, such as shared ownership or self-employed individuals. These smaller lenders often emphasise the personal touch and provide more staff time and customer support, which can be a competitive advantage over the highly automated systems of the larger players.

'Because I think this comes back to concentrating our people, our effort more with colleagues on those cases that might be even wider in the grey as such. So, if I can automate and go down more than the big banks. That's where, then the overhead comes. Yes, I've got people [as an] overhead. Let's make sure that I'm utilising where we see the biggest, I suppose, biggest return [...] I would be surprised if the market went [fully automated] in the next 5 to 10 years... because it would go against our philosophy as such that we're here for our members. If it's more automated, it's not the same personalised approach... that's why I think building societies generally say, 'We're here for our members. We need to make sure that we're supporting them. I think building societies generally would say, the banks will always say yes or no whereas we deal with the grey in between.' (Lender 5)

Barriers to new technology

Some firms developing credit risk decision-making platforms have struggled to engage mortgage lenders, finding the process time-consuming and difficult. To ensure their survival, these firms have found it advantageous to focus on other types of unsecured lending instead.

There are several challenges to fully adopting digital automation in mortgage lending. A major obstacle is that lenders continue to use outdated systems built for the credit scoring advancements of the last thirty years. These legacy systems are now large and inflexible, making them difficult to adapt to newer, more agile technologies.

'New entrants are open to using more of the new data sources than some of your historic lenders. They're not confined by some of the legacy tech and some of the legacy infrastructure.' (CR6)

One lender made an analogy between major lenders' existing systems and people who have old 'fat back' televisions which cannot access Prime or Netflix as they are not compatible with current technology. Lenders have a huge volume of data coming in from CRAs, but do not have the technical capacity to use it all. As with *Amazon Firesticks* or *Chromecast* plugins that enable old televisions to access streaming services, lenders' IT systems are currently full of patches waiting for a time when their systems can be entirely rewritten, thus slowing their ability to respond to external changes. Several lenders noted the challenges large mainstream lenders central to the market face in their ability to respond to market signals as fintech firms are capable of doing, and how much of a significant draw on the large lenders' operations these legacy IT systems currently pose.

'It's a massive issue, so that this is kind of it's it causes a split in the market in terms of the large-scale lenders that can kind of have the funding to do. Things are kind of held back by the legacy and having to kind of put all these solutions in to stitch it together, whereas you've got you kind of your newer, newer banks. That kind of have a more modern technology kit. They're able to do some of these things. I think he's there is a move to get with kind of cloud technology taking off, but also regulation and focus on resilience. So, people can be concerned about the age of technology from a security and breaches perspective moving to the kind of the more modern technologies that are safer within support. So, I think you will see that transition. But there's a huge amount of effort. Probably talking 10 to 15 years to move these banks off legacy systems onto the newer kit.' (Lender 3, consultant)

'So, we got to see the inside of how lenders operate. Day to day again, traditionally online lenders because the mortgage lenders, they, even now they're very old school, underwriting manually, different decisions depending on who put the case of the same hazards [...] I think the bigger the client, the slower they move, the less often they change their platforms. [...] Some of our more dynamic clients, they're literally changing their decision engine two or three times a day. Because they can because they will notice trends come through now. They need to respond to live traffic. The biggest of all the lenders out there Yes, they will change maybe as often as once a month. It's really rare, that they change them often. And again, there's good reason for that. But he's very traditional and then can't change your decision very often. A lot of time those decision engines are with legacy platforms. I can name a few some prior to my experience. For example, when they've used the same platform for 15 years.' (CR1)

However, lenders may not need to consider investing in new technologies if they remain competitive and profitable. Digital UK challenger banks, like Monzo, Revolut, Starling, or even Chase, set up by JP Morgan to mimic a challenger banks' approach, have made some inroads into capturing younger cohorts' current account banking, and are attracting significant funding (Cheeseman, 2024). They have done this primarily by offering superior customer experiences in onboarding, saving and payments, but are yet to compete on customer support, lending and investing (Johnson, 2021). SME lending, unsecured loans and short-term lending have been more amenable to automation and have exploited Open Banking technologies more rapidly than the mortgage market, but this may not always be the case. One consultant noted that the regulatory capital requirements for challenger digital-only banks to enter the mortgage market are a significant barrier to entry, and therefore competitive pressures from these sources may

actually be limited. Other types of lending may be more amenable to adopting new technology as there is less regulation.

Other factors that may influence the adoption of new technologies in mortgage lending are changes in the external environment from competition, regulation or elsewhere, which may also involve customer expectations, especially amongst younger cohorts. One firm reported conversations with American banks who have ambitions to enter the UK mortgage market with new technology - just hearsay for now - but other industries have been similarly shaken and such entries may force the pace of change.

'So, a lender will only really look to change if there are one of two things going on. Either they're experiencing a lot of bad debt, so their debt portfolio is growing. So that will force them to relook at what's coming in the funnel. The other side of that is if suddenly they're not growing at the same rate, they're declining more than they used to. So, it's those changes in the business model either way. So, either they're not accepting or lending as much as they wanted to do. [...] So, I would say it's less about the obstacles. It's more about their objectives and their appetite to change.' (CR6)

'There are some competitors we know that are new or existing that have huge funding, the lines, the line that I was on a call with a client just last week they're American, and they made a statement about payment. In five years time [they] will be the biggest mortgage vendor in the UK. Whether that's true or not. The traditional Lloyds, Halifax, Nationwide, they will have to innovate or die.' (CR1)

'I think it's more to do with customer expectation as well, because it's the, the speed of the service. So, as [name] was saying about the mortgages. It's, the focus is on removing any friction from the journey, so lenders are far more concerned with the, is there kind of a halt in the process and getting that decision out quickly. So, they're more likely to go for automation. Because of, because, essentially otherwise, they can't get the lending out to customers.' (Lender 3 consultant)

Other participants pointed to regulation of the mortgage market as a significant barrier to change. One firm noted that mortgage teams have access to all the same data as the lenders' credit card teams but cannot operate at the same speed of decision-making due to the regulatory oversight of the market, requiring greater scrutiny and the ability to evidence their decisions. The advent of the Consumer Duty and having to account for fair consumer outcomes, not just fair processes, may also feed reluctance to pursue AI or ML decision-making models as they have to be explainable and justifiable. However, some participants suggested that the additional data and explainable models may support lenders' justifications of decisions and outcomes and help them be more accountable.

'That's certainly some of the conversation less so on feeds of data, but more on how willing they are to go with AI type solutions because they want transparency, that ability to monitor it, understand exactly what they're doing.' (Lender 4, consultant)

'Firms need to be compliant with Consumer Duty, and one of the ways that we can monetise that is by providing an insight into the outcomes that their customers are getting right, and how can they achieve better outcomes. We can give them the data to power that. So that's another area where it can be monetised essentially.' (CR2)

But even if lenders make operational efficiencies, which may reduce their costs, this will not reduce the time to completion as other parts of the home buying process, especially the conveyancing side, have also been slow to progress. The wider array of market players involved in home buying display limited adoption of new technologies, arguably laggards in respect of

technological change. Buying a home within hours or days is technologically possible, but incentives and, therefore, motivations to change are limited due to the structure of the sector. Some participants complained that mortgages are often delayed because solicitors have not done their work, leading one lender to work on offering their own digital conveyancing arm in the future.

'So, it's more to do with the technological infrastructure of the rest of the housing market rather than the mortgage market per se. [...] Wouldn't absolve the banks of any blame here, right? I think that they're the big hitters in this. They're the ones with, you know, that are making billions of pounds of profits a year on the off the back of mortgages so they could absolutely do more to lobby and to speed up the process. But they just, they just don't need to, right? So, and I think it's unfortunate, because a bunch of reasons it's hugely frustrating to apply for a mortgage regardless of your financial situation. It's just it's the same today as it was 20 years. It was hugely frustrating.' (CR7)

'The biggest obstacle we've got to doing it... is the legal process that sits behind the mortgage. [...] The solicitors literally throw everything in at the last minute, and we tell them they can't complete their loans because they haven't done their work. [...] 'We will be doing some work ourselves next year to try and improve the process by doing it ourselves.' (Lender 6)

Fintech firms have expressed frustration with the slow pace at which the mortgage industry is realising its technological potential. Although the previous government conducted an inquiry into digitising and improving the home buying and selling process, with a focus on using smart data and digital property packs to speed up conveyancing (Hollinrake, 2024), there are concerns about whether rapid home purchases are desirable. Homeownership can come with significant challenges, and thorough due diligence is essential for making sustainable decisions (Wallace, 2016; Kear, 2017). Therefore, it is crucial to balance speed with obtaining proper advice and guidance. However, while regulation of various actors in the home buying process is necessary to protect people and markets, there is undoubtedly still room for algorithms to enhance the mortgage lending process and create efficiencies.

Conclusion

Much of mortgage lending is already highly automated. This chapter discussed the next phase of digital transformation in mortgage lending, focusing on making customer journeys smoother and automating administrative processes that currently rely on manual workarounds. Technology firms are advocating for the broader adoption of new platforms using AI and ML, which have demonstrated in other lending markets their ability to be more agile, reduce losses, and lower staff time and costs in administration and decision-making. However, the adoption of these innovations in mortgage lending faces barriers, including outdated legacy systems, concerns about regulation and accountability, and the ongoing importance - and sometimes benefits - of retaining human intervention in the process.

Chapter 4: Data possibilities

Introduction

This chapter explores the project findings relating to the range of data sources available to lenders, from traditional sources like CRAs, credit bureaus and credit information brokers, to newer forms of data obtained directly from various suppliers via APIs. Traditional credit data, such as credit history, is often enhanced with official statistics like household income and expenses from the Office for National Statistics (ONS). Additionally, newer data sources include tax records, wage slips and banking data. These diverse data points, numbering in the hundreds for each credit applicant, have grown significantly in recent years due to advancements in IT systems capable of managing large volumes of data. This expansion not only speeds up verification processes and enables automation but also provides deeper insights that enhance lending decisions.

Open Banking transaction data offers quick and detailed assessments of an individual's income and spending habits, providing a comprehensive view of their financial behaviours, strengths, and weaknesses, thus reducing time and costs for lenders. This data holds significant potential for mortgage credit risk decisions through algorithmic lending processes. However, similar insights can also be obtained from data digitally extracted from PDF bank statements, which bypasses some regulatory constraints associated with Open Banking. While other types of digital footprint data can estimate a person's likelihood to repay loans, their use in the UK credit industry is limited, confined mainly to fraud detection and flagging applications for manual review, with little adoption beyond these areas.

The chapter explores these data sources, starting with traditional credit data, then it examines the expanding role of banking transaction data and the new insights into financial behaviours that inform credit risk decisions. It also covers digital data related to identity, followed by other emerging data sources that are currently available and in use.

Credit data

Lenders evaluate loans using three main data sources. To assess an individual's credit risk, they combine (1) information provided by the customer in their application with (2) data obtained from external sources like CRAs and other direct suppliers. Additionally, (3) a property valuation is conducted to ensure that if the borrower defaults, the lender can recover their funds by selling the repossessed property. This valuation is carried out by surveyors or through digital tools known as automated valuation models (AVMs). Each of these data points may seem insignificant on its own, but when combined in predictive models, they become highly valuable in forecasting outcomes. This chapter explores the data obtained from CRAs, including traditional credit history and newer data types that CRAs or credit information brokers now offer. These combined data sources are then used to evaluate the customer's creditworthiness.

The lending industry has experienced relatively stable technological operations for a long time, with the main exceptions being the global financial crisis and subsequent regulatory changes. Credit scoring systems derived from a range of credit market information were adopted in the UK from the 1990s and are the mainstay of the industry's lending decision making processes.

'There is a role for a credit score in the sense of an algorithm so put in some data, and you get a score that indicates a level of credit risk worthiness that we also do. That is an unambiguous yes, for the major providers for everybody. They will always calculate a credit score. [...] The manner in which they calculate these credit scores...practically hasn't really changed that much in the last 25 years.' (Lender 4, consultant)

'We introduced credit scoring in 1996/97. We, we had automated credit scores, but we didn't throw away manual assessments. So, we used to, we used to use scoring as a first line of defence, if you like, on our credit analysis. Right. But we would routinely sample the credit score decisions to make sure that we thought the score was doing the right thing and we would routinely. So it comes out with a sort of red, amber, green light. These ones are the score says no brainer. Do it. The next one is, it's failed, but we set a tolerance as to it, only just failed. And then there were the ones that were red. Well, we'd sample all of those to determine, you know, was the score doing the right thing. And, and we used to set our own calibration on where the red, amber green was. In fact, that is what happens in lenders. That's what happens now. Yeah. That was new then though.' (Lender 7)

The marketing of credit scores and encouraging consumers to check them can be somewhat misleading, as the specific score shown to consumers is not directly used in lending decisions. Credit scores vary between CRAs, and lenders typically access a range of data from CRAs to build their scoring models. These custom scorecards apply unique weights and reflect the lender's specific risk appetite and target market segments. Therefore, lenders do not use the credit score numbers provided by apps like *Credit Karma* or directly from CRAs; those scores are primarily for marketing purposes.

'The credit reference data we ingest is the basic raw data from a credit reference agency. We import a queue of information. From the credit organisation, including a score application score. But we don't use that because there's no correlation between application score, applicants' score and performance. So yeah, so what we do ingest, is we ingest all the data related to any credit cost performance within their credit cards, mortgages, any other outstanding.' (Lender 6)

Lenders are increasingly inclined to develop their proprietary scoring models rather than rely on industry-wide models. They do, however, heavily depend on credit information from CRAs or brokers, which typically includes adverse credit markers (such as County Court Judgments (CCJs), insolvency and bankruptcy data), a 13-month rolling history of payment performance on outstanding credit balances and electoral roll data. Some lenders pull data from multiple CRAs, noting that these data resources are continually expanding.

A secondary industry has emerged around the acquisition and resale of data within financial services and beyond (FCA, 2023). Lenders combine these external data points with their own customer performance data to create tailored scorecards that align with their specific business requirements.

'Most of, I suppose, different bureaus hold different information... That's why we tend to go into one or two bureaus instead of just relying on information from one source. [...] We're checking for accuracy and also adverse credit... we're starting to see more information... so it's building a wider picture on what that data might look like. [...] Now, some of the data bureaus are getting much, much more, we get...we're seeing more information. Therefore, you can start to look at that. Your gas bills, you travel car insurance, etc., that's all starting to come through. So, it's building a wider picture on what that data might look like.' (Lender 5)

'What I would also say is that ten years ago lenders would use an off-the-shelf model and credit score that maybe the credit bureaus [CRAs] provided. Most lenders nowadays are opting to take the raw, underlying data and build their scorecards and models, because they can train it on the

historical performance of their customers. So, a credit bureau will build a model based on the general population, whereas if you're a bank you will build your model on the historic nature of the customers that you've had. So, it becomes far more fine-tuned and specific to the demographic and type of customer that you tend to attract as opposed to it being a generic, UK population scorecard.' (CR6)

'So that's been we've been doing that for decades of reselling [CRA] data. And two years ago now, we also did basically exactly the same deal with [another firm], where we retail experience data. So, what that means to us is our client can have one contract with us for the data products, the classic process AML [anti-money laundering], bank check and affordability check those forecasts of products. Most decisioning may change. We can supply that from [various CRAs].'
(CR1)

This expansive pool of data that the market makes available brings further issues for lenders of aggregating and reconciling often error-prone or conflicting data. The FCA (2022) credit market information study notes that there are more credit files than the adult population, as some files are duplicated, with missing records on some files and duplicated entries on others. If lenders are drawing data from a single CRA, likely, they are also only reporting debts and defaults to a single CRA so the picture can be incomplete. Larger mortgage lenders were more likely to look across the bureaus for indicators of risk, but how they treat or identify multiple missing or duplicate data is a problem. Lenders seemed sanguine about the possibilities of incorrect credit risk and affordability assessments arising from these duplications, errors and omissions. Some lenders use the credit information data in rudimentary ways, acknowledging the errors and difficulties in matching data. Other firms are producing software to 'de-duplicate' and help improve data quality and increase its utility, supporting lenders to maximise the use of the vast data reserves for their specific business needs. Other participants recognise the CRAs have substantial expertise in this arena and are improving data resources and quality, but that the speed of innovation and lending in some sectors means that CRAs and regulation cannot catch up. For instance, instant short-term or other loans are available but credit reporting is done only monthly, so information asymmetries remain as there are no real-time records that reflect a person's indebtedness or gearing, and nothing to prevent them from therefore making additional loans before the next reporting cycle.

'What we do sometimes get revealed is sometimes, so we know [for] online lenders, a credit score that's based on the telecoms industry, is actually quite predictive. [That] wouldn't work so much for mortgage lending. So, there's different flavours. Lots of different data sources go into these credit scores, all different compositions. It's not just credit files, it's loads of things. It's the how many credit searches you've had registered, it's your electoral roll. It's other things in fact that facts, bring out some credit scores. I believe that based on combining Open Banking data into it as well. So, there's a huge amount of different varying sources and some credit scores work well for some divisions of lending and not too much for others.' (CR1)

'It's a known thing which is different from saying it's, I mean the quality of data that you get is always that circumspect and you're always very careful with it. The credit bureaus will say the banks rely very heavily on this, and they rely on our analytics, everything else, and that's complete bullshit... It's useful information. But they're well aware that there's a lot of crap in there as well, and they tend to use the basics of it. So, I know one that is just literally all they care about from experience is, do you agree that this person is a person and this, this address? They're not interested in much more than that.' (Lender 4, consultant)

All this data helps appraise creditworthiness and is used to indicate the accuracy of consumer-supplied financial and housing histories relayed on forms. Lenders will take data on employers and employment tenure at face value and check it against the data from the credit bureau. The credit history data is used in different ways with small sums of adverse credit often

discounted, while other lenders may exclude people with any red flags, such as gambling. Credit histories and CCJs' weight depends on their value, timing and importance, but are a critical component of lenders' behavioural propensity to pay scores.

'For example, some lenders won't touch you if you've used a payday loan in the last 12 months. Some lenders don't like it if you've got any gambling coming in and out of your bank account, so quite specific. In general, the general thing is there's a hierarchy of the types of payment, types of missed payments. If you have a missed payment for your telephone bill or for a parking ticket, which is less than £500, a lot of the lenders that we work with, they will turn a blind eye if you've only had one in the last 12 months. If you've had a County Court judgement, so it's obviously taken a little step further, a lot of lenders again will turn a blind eye if there's only been one of those in the last two years. Again, it depends on the size. So, £500 seems to be the figure that people look at. If you've got a £5,000 one, then people are going to be starting to get a bit more worried. If you've then got a missed mortgage payment, people, again, will be getting a little bit more worried. Then, obviously, if you're going into some sort of IVA or if you've been into bankruptcy, then in reality, you're unlikely to be able to buy with any decent rate for about three to five years. The credit score is almost like a marketing tool.' (CR3)

Some consumers are skilled at managing their credit scores, understanding them as a summary of key financial data. Several interviewees tracked their scores using apps like *Credit Karma*, and one successfully improved her score over time. This focus stemmed from past financial ties with an ex-partner that had negatively impacted her creditworthiness. She improved her score by paying off credit card balances in full and taking out small, manageable credit lines, such as a card used only for groceries. Borrowers were aware of the importance of presenting themselves well to financial services. As noted at the outset of this report, this aligns with Fourcade and Healy's (2024) concept of '*eigencapital*,' or digital social capital derived from online activities, highlighting the growing pressure on people to manage their digital profiles.

Banking data

Open Banking technology holds the potential to recast how income and affordability are verified and assessed. If anything encapsulates Fourcade and Healy's (2017; 2024) comment about algorithmic profiling shifting approaches from firms on the *outside looking in* to now being on the *inside looking around*, it is Open Banking. No longer looking just at our payment history records and demographic data, credit inferences are made on our whole financial, and indeed other behaviours. Accurate and almost real-time income and expenditures can be determined, as well as the velocity of spending and our spending categories and patterns exposed for judgement. This is intimate data, although borrowers have already, especially after the financial crisis and tighter mortgage market regulation, had to supply bank statements to support detailed affordability assessments. Arguably, however, Open Banking technology is automating and maximising what can be already gleaned from paper or PDF bank statements anyway.

Although, lenders and firms have developed software that 'reads' PDFs or JPEGs of bank statements, much like recent payment technology 'reads' data from bank cards or cheques, entering the relevant information and figures into their systems and classifying spend in similar ways, bypassing the tight Open Banking regulatory framework. For these reasons, some participants, particularly on the lending side, were sceptical about any transformational potential but the speed and analytic insight afforded by Open Banking could be profound. While financial education tells consumers how to improve credit scores, keeping balances below certain proportions of their credit limit, opening credit cards or taking subscriptions to provide data in

otherwise thin credit files, this guidance is otherwise silent on Open Banking and associated technologies and their analysis of our day to day spending financial behaviours.

Open Banking data was pushed by the UK government as a tool to increase competition in the financial services sector. It allows other firms access to the customer data insights that only current account banks hold, which would, in theory, enable customers to get better deals and products from a wider range of firms. There is also a provision to remove third parties from payment systems, but it is on the banking transaction data that innovation has had the most impact (Kassab and Laplante, 2022). It is designed to give consumers control over their data, who they choose to share this information with, providing access to other services and tools for money management (Reynolds, 2021). One technology firm outlines its potential from a consumer perspective.

'When you think about the point of Open Banking, like why was Open Banking brought in in the first place, well, it's about increasing fair competition in the market. What it does is it gives me as a consumer, just not [Platform company] me, but me outside of work as an end consumer, it gives me the power over that data to say okay, well I bank with Santander, but at any point in time I can say give my data immediately to this third party, and they have to comply with it. Which in my mind kind of makes me the owner of my data and not Santander. They are a custodian of that data, but I'm the owner of it. That's how I feel as a consumer, and that's what Open Banking I think has done for me. It allows me to do things like, I use an app called *Sprive* which is a fantastic app out there in the mortgage space actually, that looks at my affordability data, and my Open Banking data every week, and it sets aside for me a certain, what it thinks I can afford, and it sets it aside for me to make an overpayment on my mortgage, for example.' (CR2)

Open Banking and alternative credit assessments were discussed as tools that could enhance the accuracy of credit scoring, particularly for individuals with limited credit histories. This accuracy is crucial, as borrower estimates of expenses can often be imprecise; for example, one borrower simply multiplied their daily spending by 30 to estimate monthly expenses, neglecting irregular or periodic costs. Open Banking is central to short-term lenders' operations and helps to provide loans to individuals with adverse credit who have no other access to finance. Another lender uses rental data provided by Open Banking technology firms to support mortgage lending decisions. The technology is framed as a key tool to overcome the problems of thin credit files. One firm suggested that this approach aligns with the Consumer Duty regulation (FCA, 2022a), which requires lenders to offer products that aim to achieve better customer outcomes, not just maximise shareholder value.

Open Banking provides precise and timely affordability assessments and, by categorising transactions, offers insights into the borrower's life, helping lenders tailor the level of service or support needed. Open Banking is also said to reduce lender time and costs, delivering more consistent results than manually reviewing income and expenses via bank statements. Many firms believe that Open Banking technology provides a more accurate view of a person's ability to pay, contributing to a fairer financial system. However, its use in the mortgage market remains limited.

'So, I think new entrants to the market. So younger people or students or people, even people that have been renting that are coming into buying properties. It's still really, really difficult. And I think it's my view on it is that I think it's due to a lack of adoption of the technology that's available to banks or financial services, organisations to make better decisions. So, we talk about Open Banking and transactional data now being available, not a majority of banks that [are] using it. And I'm not using it, they're certainly not using it in mortgage applications. It is sufficient to be able to address those more marginally, more marginalised under thin credit file individuals and

thin credit files [are] tricky, because everything that the bank does, everything that they do from automation depends on data. So, it depends on being able to not just take you at face value. But...it's almost like the Matrix, right? They look at you and see a scroll of green data. And that's where they're gonna make the decision. If the screen is blank when they look at you. It's very, very difficult for them to automate that.' (CR7)

'As well as regular income transactions, we can get regular outgoing transactions as well. So, we can see from the data where a client is consistently paying their rent on time. That is a, for those kind of credit invisibles, that's a really important point to be able to demonstrate, "Well okay, I haven't got ten credit cards with loads of repayment history on it. So, my credit score maybe necessarily isn't that high, but essentially, I'm a renter, and I now want a mortgage. I can demonstrate through my Open Banking data that I've paid my rent on time every month for the last 12 months", which is, or should be, an important consideration in those decisions.' (CR2)

'It's a fairer picture of what the potential client can or cannot afford [...] So for an underwriter it's incredibly valuable to have that same picture when someone's applying for a loan. Historically when you look back at underwriting practices, it's a case of okay, well let's, and I don't know if it's the same for you, but when I went to get my first mortgage, I had to take with me a bunch of printed out bank statements. My mortgage broker sat there with ten different colour highlighter pens, going through it for hours and hours on end. So that's the problem that we're trying to solve. How do we make that ten-hour process into like a ten-second process?' (CR2)

'Quite startlingly, we did a retrospective analysis with TransUnion back in December of last year and we took 30,000-odd customers - that's customers and not customers - and we ran them through the TransUnion bureau, and we have lent to ten percent of our population who don't even have a credit footprint with TransUnion.' (CR4)

One consultant saw Open Banking as offering more granular data, but the turnover and detailed transactional data are available from credit bureaus and borrowers' own banks anyway, so considered the benefits of Open Banking to be overstated. Not least as CRAs are getting more sophisticated with the data they capture and summary analysis of turnover data they provide, and are getting involved in Open Banking themselves, so lenders may not have to engage with these specialist third-party services directly. Another consultant was sceptical that the new data changed anything fundamentally. Open Banking might offer more insights into customer behaviour, but one lender was sceptical that it does not necessarily transform the existing process. The professional resistance or at least scepticism on display echoes sentiment in the literature about the shift from manual underwriting to the use of credit scores and automated lending in an earlier period (Lauer, 2017).

'Bluntly speaking, the banks can get that data in an aggregated format through the credit bureaus, and they'll get it on a monthly basis. So, it's more just an ease of having more granular data. But they can get on it. And the amount of data. And typically, if you, if you bank with a bank and they've got your current account, they have this data, they have your current card. So, there is a perception more than a reality that needs to be resolved.' (Lender 3 consultant)

'I didn't see the advantage or disadvantage again. It's just get the paper statement for that. It makes it easier to process it. Put it through an algorithm. It doesn't fundamentally change what we're going to do.' (Lender 4, consultant)

Open Banking technology is still in its early stages, and several challenges have been identified: handling multiple accounts and bank transfers, transaction categorisation, and high consumer dropout rates. One short-term lender using Open Banking noticed that borrowers frequently moved money between accounts, raising concerns about whether the technology can detect such behaviours or if borrowers might game the system. This has led some lenders to believe

that widespread adoption of Open Banking may be far off. Technology firms are aware of these concerns and have developed solutions, such as analysing long-term consistency in incoming funds and managing multiple accounts, including those from the UK and Ireland, to improve the reliability and accuracy of Open Banking assessments.

'So it might be that my mum has sent me £500 to help me pay for something, and that won't appear as income, that'll appear as an incoming transfer. So, we kind of draw a line in the sand there. I think different Open Banking providers will have different levels of sophistication with that, but for us it's like, well if we can't prove the consistency and the longevity of it, we're going to pass it as an incoming transfer, so money that has come from elsewhere that is not a salary, or regular income.' (CR2)

'It's the granularity of the data. So, you did so your salary did come in, but as your salary moved? Is your salary flexed up or down, what's your average spend? What's your average spend, and your maximum spend, and your minimum spend? What's your average? What's your number of direct debits? And is that flexed up or down? What's your payment? History everywhere else? Are we seeing regular payment payments come in from other financial services organisations that you deal with. Have you changed, employer? If has that been flagged through? Have you missed any transactional payments is your payment as your as your disposable income flexed up or down over a period of time, and all of that is on a, as I say, we that's taken. That view is taken monthly and then it's rolled over 13-month period, so that we get a consistent view of behaviour.' (CR7)

Open Banking technologies provide profiles of a customer's income and expenditure, identifying income from benefits or earnings, and categories of expenditure such as food and utilities and can identify 'red flag' categories like gambling. These insights into income and spending patterns over time can be undertaken rapidly. Some insights may confirm customer information, others reveal things that would influence the affordability assessments like undisclosed borrowing, child support or support decisions about the significance of gambling to any underwriting decisions. Gambling does not show up in CRA data and for one short-term fintech lender comprised an important spending component for 29% of their borrowers. Other red flag events were continual use of payday lenders, overdrafts, returned direct debits and buying cryptocurrencies. One borrower reported that the mortgage broker had steered them towards lenders that scrutinised these issues less.

'For instance, one of the big things, it always makes people laugh - the amount that people spend on *OnlyFans* [pornography site], for instance. It [Open Banking] is the absolute unbarring of most people's souls. Then, for instance, a big new craze that seems to have hit everyone is raffle tickets. Win a house, a luxury house. There are people spending hundreds of pounds a week on that. Again, is that gambling? We think it is. It's classed as gambling but, again, that's a non-reported thing. Of course, it's in people's own banking... Cryptocurrency. If that's not bloody gambling, I don't know what isn't. We've had people going, 'I'm an investor.' Well, why do you need to borrow £1,000, then?' (CR4, short term lender)

'Hopefully when they start regulating crypto and buy now, pay later - a lot of lenders won't accept deposits that have been got from crypto, so that's a blocker. People have to go and put it into a savings account and leave it for six months, so it's been cleaned up, but they won't accept crypto. I think what happens around the data and approach to crypto and buy now, pay later will have a big impact. I think those would be my areas where data will improve our decisioning.' (CR3)

Spending categorisation remains a challenge. One firm, using technology to analyse bank statements instead of Open Banking data, categorises income and expenses according to ONS categories such as utilities, food, clothing and gambling/alcohol. They compare these categories with average household spending in different locations, recognising that costs like food or

entertainment vary between cities like London and Leeds. The aim is to isolate essential expenses for affordability assessments, as required by FCA mortgage regulations, which mandate fair treatment of customers and the use of essential, not discretionary, spending in loan affordability evaluations.

Lenders have concerns about the accuracy of Open Banking firms' categorisations, especially when distinguishing between essential and non-essential spending at various outlets. Food staples may be purchased from Amazon or electronic goods from Sainsbury's, confounding simple classification. There is no industry standard for these, and opinions vary on whether the regulator should step in; however, given that no universal standard exists in credit reporting after 30 years, some believe regulatory involvement is unlikely. One firm noted that categorisation accuracy is improving through consumer feedback and ML. Another mentioned that when their technology, which reads PDFs of bank statements, wage slips, or tax returns, cannot classify certain data, those cases are referred for manual underwriting. Additionally, one firm pointed out that bank statement data is often used to confirm a customer's address and business location, which Open Banking data currently cannot verify.

'Personally, I think Open Banking has got, it's got a number of use cases. But to get it there for kind of the affordability and the credit assessment. I don't know. I don't really know how you would ever get it down to that kind of essential spend category. And you're gonna be, gonna be more profiling on customers' behaviours. And you might use that, as this customer puts a lot of money into savings on a regular basis, therefore they're a positive. I think you can get inferences from that. But it's not gonna replace the need to capture income and expenditure data? Right?' (Lender 3, consultant)

'Banks used to get bank statements are looking for 3 things, are they a real person, does the bank statement suggest they are who they say they are, does the salary accord with what they've told us, do they have that income; and lastly what they spend money on, banks don't like gambling but do like baby things, for example. But AI can do the last one in a way that humans find it harder to identify as they have to go through too many lines of data but not infallible. An open statement will tell you that you spend 50 quid at Tesco. It wouldn't tell you whether you spent 50 quid on Pampers or 50 quid on cider? Right? Yeah, it's not gonna tell you that.' (Lender 4, consultant)

'Wherever they tell us, it's [the categorisation of data] wrong, we store that data, see what they're re-categorise it to, and that data gives us the opportunity to go back and retrain our machine learning model, to make sure those insights, and that accuracy is always recycled and put back in. So, it's always getting better all the time because we've got that active user base through the app.' (CR2)

These Open Banking conversations also highlight how involved people are in algorithmic classification, not just the code.

'There is no standard around the classification of Open Banking data. [...] I don't think it's a concern, because ultimately, it's just an interpretation of the raw underlying data. [...] Yes, so even if you were to standardise the classifications, the lender would interpret the classifications differently.' (CR6)

Consumer views on Open Banking varied widely. Some borrowers were already using Open Banking apps to aggregate their accounts and find the best deals. Others felt it was intrusive when lenders required access to their accounts for mortgage purposes, especially when asked to explain specific transactions. One firm noted that younger people tend to be more accepting of Open Banking technology. However, this posed a challenge for a lender using Open Banking

data to consider parental income and assets for first-time buyer loans, as it deterred older parents who were reluctant to share detailed financial data.

Lenders have also observed public hesitancy in granting consent for Open Banking, largely due to trust issues and perceived risks associated with the technology. Although Open Banking is promoted as a tool that empowers consumers by allowing them to use their data to secure better deals, in reality, there are often no clear incentives, like better rates, for providing this additional insight. One lender tried to reassure customers that they were only examining aggregate figures, but others did delve into detailed, granular transaction data, leading to varying levels of comfort and acceptance among consumers.

'So, we had Open Banking initially, and a lot of our work is around families and friends, so have slightly odd, a lot of our key decision-makers are in their fifties. Not that that has a huge impact, but what we found was Open Banking is still a huge blocker to people working with you. People don't want to hand over their bank's details, particularly it's okay if you just hand over a small bank, which is not, if people have multiple bank accounts, if you just hand over the banking details for one that's got £500 in it, whatever. If you've got all your life savings, people are really, don't want to hand over those details for Open Banking, and so we just found people just dropped out...' (CR3)

'I've got one called *Snoop* that tracks my bank accounts... I've given them permission to do that, and I've saved money with them in the past, like changing mobile contracts and stuff, and insurance companies. I was a little bit concerned at first... It was a bit nerve-wracking doing it, but I was fine. It was a necessary evil that I had to share my bank details with somebody. [...] I don't like people having – I'm all right with giving people my data, but I don't like these people who sell your information on. I don't agree with it, I don't think it's fair, I don't think it's right that they do it.' (MB4)

'I had to go through every £ and try and explain where it had come from, so all my bank accounts were linked to this *Armalytix*, and they had to... I don't know how it works but obviously they assess it all, take all the data and then provide a report, but it was very difficult to actually complete as well. It took me a long time because they were asking about, where did this come from, how did you obtain it? It was just, I don't know, it just felt like a lot of work, and it was very, quite difficult to be able to explain where everything has come from over the past ten years I've been trying to save up for a deposit.' (MB1)

In the private and social rented sectors, tenants raised issues of consent around banking data being used to access housing. Perhaps there was a greater expectation of comprehensive data sharing when applying for a mortgage, but borrowers were more accepting than tenants. Part of that may be about trusting banks with data security. There is also greater regulation and scrutiny of financial services in terms of the assessments and outcomes achieved than with landlords and letting agents.

'I think that the big part the regulator has to play in data is around the security of it. So, making sure that people only have access to the data they should have access to and that they hold it securely and that kind of stuff. In terms of what they use that data for, then that's regulated in the sense of the permissions and the explicit consent. The challenge with that - and you mentioned it earlier - is that if the consumer feels that they're over a barrel and therefore feel like they must provide this whether they're consenting or not, if they believe they have to then I think we're at the wrong end. Consumers aren't there making an informed choice as to whether it's the right thing for them to share Open Banking data.' (CR6)

Firms were also aware that customers made trade-offs, setting the unease or risk of sharing data against the value of what they might obtain if they do so. Obtaining a mortgage and buying

a home was therefore a strong imperative for people to give up data. Reportedly, conversion rates for app-based banks were high, typically over 80%, but for other banks and their customers as low as 50%, whereas credit data achieved a higher rate of success, explaining lender reticence to rely on Open Banking and use it in supplementary situations only. This conversion rate mapped age as noted, but also banks' own technology, how smooth, reliable and reassuring their interface to the third-party Open Banking provider was. One lender did not want to interfere with the brokers' relationship with the borrower by introducing obstacles and third-party apps to the originating process. Firms acknowledge that Open Banking has faced some resistance due to concerns about data privacy and security but consider that adoption could be accelerated by building trust and increasing consumer awareness.

'There was a big study done, and it was about data ownership, and it was around the, it was actually the mindset of a consumer, and how if I'm going to give you my data, what is the trade-off that I'm going to get for it. So basically, the value, probably again without being ageist kind of thing, but the value younger consumers are putting on their data, in that they're not going to just hand it over. "What am I going to get as a response to me giving your data, where is the value I'm going to get as a response to that.?" [...] That trust has to organically grow, and the way that it will organically grow is by driving good outcomes, and people getting better loans, that they either couldn't get before, or that are more affordable to them. Oh, okay well, I used my Open Banking data here, and actually I ended up with a much better deal, so in the future I'll be more likely to trust it. The more of those stories that permeate the public consciousness, the greater trust will grow.' (CR2)

'The take-up rate on these things is really low, people's willingness to do it is just really low. {the purpose] was to really transfer the response, the kind of the ownership of the data back to the customer. And with a view that you owned your data, you could decide what it was worth and the value. But I don't think anyone's really monetised it correctly to go right if you, if you let me have access to your Open Banking. I'll give you a reduced rate on your thing, or what something like that that get an incentive.' (Lender 3 consultant)

'It's a lot of data to give away. I can see why people would be dubious about it because it's a lot of financial information. [...] I think as Open Banking gets more and more use cases, and is more and more common, then that trust will build up.' (CR2)

'We're not talking about getting a store card. We're talking about getting the keys to your property. Therefore, consumers are more willing to maybe step through a few more steps and allow companies access to more of the personal data that they are a bit more nervous about sharing. So, I think there is a huge piece for the industry to get around as to that value exchange. So, what does the consumer get by allowing access to these additional datasets? That's a key thing to remember in that whilst some of these datasets are publicly available, the holy grail for some of these underwriters is those that they don't have access to, and the consumer must explicitly give them consent to have access to.' (CR6)

Participants closer to the mortgage industry tend to be more sceptical about Open Banking compared to technology firms or unsecured lenders. For instance, some larger lenders have opted not to accept Open Banking assessments as sufficient proof of income and affordability, preferring to use PDF bank statements instead, leading them to discontinue Open Banking technology. A credit lender with mortgage experience believed that the mortgage market is lagging and could benefit from Open Banking to expand homeownership access, given that mortgages are secured by property, which mitigates risk. Participants saw more potential for Open Banking in areas beyond mortgage lending, such as providing real-time insights for early intervention with struggling customers or targeting those with improving financial situations for product marketing. Open Banking's ability to track consented banking data beyond the initial

retrieval also offers opportunities for post-loan assessments and evaluating the social impacts of short-term loans. While Open Banking applications in mortgage lending may currently be limited, they promise to enhance customer support, loan management, and evidence for the Consumer Duty.

'What we know Open Banking is, in a simplistic sense, is, 'Can you give me your current account history for the last x out of many years?' Which people have been doing by providing bank statements, come forever. So, it's a slicker way of doing that. Yeah, it's not exactly revolutionary and because people have been doing it for other, particularly in the mortgage market, they've been doing it manually, but they've been doing it forever. it's not going to make a big difference. '(Lender 4)

'Because they're pretty archaic! A lot of the banks have got very archaic processes, systems. Changing them is quite a big thing. They want evidence in traditional forms. They want PDFs. Yes, you can do all the Open Banking, and then you still need to ask for the bank statements anyway, so it's like, well, there's no point. It's like what are the friction points? I'm really not a big fan of Open Banking because I think it's great for the little apps for information, but actually, when you go to mortgages, big amounts of money, is it really that big a deal to go on to your bank statement and download and send it over?' (CR3)

'Open Banking and transactional data really allows banks, when used properly, to massively transform and get better quality decisions based on what they're seeing for you, actually, as to what happened yesterday, or even this morning, not having to wait from trended monthly rollover of big ticket risk things that you've been doing like making payments or not making payments. But are we seeing the fact that as a salary not gone in yesterday, and it should have. So, what's happened there? And are we seeing the fact that you've you? You did miss your mortgage payment on Tuesday. [...] So we're looking at kind of almost what we call early delinquency measures. [...] But equally the lenders will use it candidly. They'll use it to cross sell, and upsell, so we'll be seeing an uptick in behaviour. So [someone] has met a mortgage payment. She's just paid a loan off actually in full. She's overpaid on a credit card. A salary's come in. It looks like it's 20% more than it was. Okay. Should we now be thinking about reaching out to her? [...] Does she need a, should we be selling her a new loan, a credit card, a mortgage so, really, really rapid.' (CR7)

'I worked at [social enterprise] and we pioneered an I&E [income and expenditure] statement using Open Banking because a lot of people when they go and look for debt advice - by the time you go looking for debt advice, you're already in problems. At that point, you've lost touch with the reality of your debtors and creditors. This was a great way, we felt, of bringing that all to light, so they got everything in there. I do think it's still providing us with new and exciting areas to look at in terms of what patterns we have; both from a non-payer profile through to a paying profile and whether or not it's right that a customer that comes back to you having shown good payment processes should be entitled to a lower APR [annual percentage rate], for instance.' (CR4)

Identity data

Digital identity verification tools are widely used in the housing market. Lenders employ systems that verify the authenticity of personal information and documents, and some include facial recognition software. For example, one system requires borrowers to take a photograph and record a video to ensure they are not being coerced and to validate passport data. Background checks can identify politically exposed individuals and support anti-money laundering compliance.

One firm discovered errors in passport data, which were found to be mistakes at the Passport Office rather than with the borrower. Identity checks also verify immigration status, a growing

concern for lenders who need borrowers to have been residents for at least three years and have at least 12 months remaining on their visa. This is increasingly relevant for professionals coming to the UK for roles in the health service or engineering who wish to buy a home. The issue is expected to grow as the settled status scheme for EU nationals post-Brexit allows five years to formalise their immigration status, and many may not have completed this process yet.

'It confirms to us that the information they give us in terms of their phone numbers are correct. They're alive from a fraud risk perspective.' (Lender 6)

'Yes, so we use *Onfido*, so that's basically so people will when they've done the plan, they get sent a link and they take a photo of themselves holding up their driving licence or their passport, and that then validates them, who they are. Yes, *Onfido* is really good.' (CR3)

Other data sources

CRAs and data brokers now provide mortgage lenders with a broader range of data. However, lenders are increasingly turning to other direct data sources. They can access information directly from Companies House for details on company directors and business accounts. Additionally, lenders are exploring ways to automatically integrate data from the Land Registry, HMRC, electronic wage slips and DWP benefits. Some lenders prefer obtaining data directly from these key services via APIs rather than relying on CRAs or brokers, who aggregate large amounts of population-wide data.

'Now, some of the data bureaus are getting much, much more we get, we're seeing more information. Therefore, you can start to look at that. Your gas bills, you travel car insurance, etc., that's all starting to come through. So, it's building a wider picture on what that data might look like.' (Lender 5)

'Personally, I think if I was the government, I would mandate banks used HMRC data as the income, because that would help the government collect tax because it means any sort of tax avoidance meant you couldn't actually get any lending, so you would kind of you would close 2 loopholes by mandating that. And, secondly, it's the most trusted source of income that you can find. So it's, it's validated. It's checked, and it would mean that you were really certain the amount that that customer actually earned. So, from a responsible lending perspective, you can really lend and make solid decisions based on that information. Now, people that have obscure incomes or don't have income that's registered through tax will lose out or not be kind of beneficially treated, but from a control and a protecting consumer, I think it would be a better way forward.' (Lender 3, consultant)

'I'd be very worried about people like Experian getting their hands on more personal data because all they want to do is sell products and build, target products more. I hope they don't get their hands on things like student data [Suggested in Experian (2018) Invisibles report] because it seems a bit immoral to me, but I'm - anyway. I think the way people can access data is better. We don't get generic data because we are targeted, we're focused on a specific, we focus on specific individuals, so they provide us the data to them, about themselves. Whereas Experian is trying to build a big data set and they can target groups of people, we're not doing that. We're looking at the individual when they come to us, and we get the data that we need and then we validate.' (CR3)

Spatial data was once used in lending decisions via neighbourhood classification systems like ACORN or MOSAIC (Webber and Burrows, 2018). Still, several participants indicated that this was used less, and certainly not in formal decision-making. These data reflect the attributes of neighbourhoods and not individuals.

'I would say they're less and less common these days because they're just not as powerful when you're doing kind of your analysis that just don't come to the top of the list of the variables that you want to pick.' (Lender 3, consultant)

Residential stability has long been associated with creditworthiness, but as people with otherwise good financial profiles are staying longer in the private rented sector with limited security of tenure some lenders have adopted other measures of steady, orderly or consistent lives.

'One of the things that is becoming a little bit of a proxy for stability and security is how long people have had their mobile phone numbers as an example. If you're like me, I've had the same phone number for 18 years and it never changes. Certain datasets out there can tell you how long that individual has held that phone number and how long that phone number has been active. Now, if you've got a mobile number that has only been active for 30 days, that's going to give you an inference of, a) how stable is this person from a financial lending perspective, because if people are swapping and changing their mobiles, well, b) why are they doing that? So, is this because they've got credit people chasing them, so they get rid of that? Similarly, with email addresses, whether email addresses are valid and how long they've held them are useful proxies in terms of how stable people are.' (CR6)

Real-time data and learning models that can change swiftly may be appropriate when external circumstances change. One firm noted the limited utility of pre-pandemic models to post-pandemic spending as patterns had altered significantly. Likewise, interest rate rises have switched the benefits of borrowing and savings, changing people's spending and saving patterns.

'I think the key bit from a technology perspective for me is how quickly we are now required to adapt and change. So, something that interested me through COVID, a lot of the traditional models that a credit bureau had built up... So, a credit bureau builds models based on historical data. So, it looks at trends and then uses the trends of the historical data to predict what's likely to happen. Now, from a credit perspective, spending behaviour changed significantly during COVID. People weren't travelling so much. They were spending money in different places. You couldn't buy a new car, so people weren't taking car finance. People were tending to do more home improvements and to save. So, where you've got two years of people's finances behaving differently, the models pre-COVID are no longer very predictive.' (CR6)

Lenders often use ONS data to capture or benchmark household expenditure, but this lost some utility due to it being out of date, a problem that new 6-monthly data releases have not overcome. Some lenders, not necessarily mortgage lenders, intimated that they were no longer using this data source.

'They use ONS data in all their models and things you can draw on the back. That's a whole heap of data available through ONS. So, what sort of typically might you be drawing on there is this house price, income wealth, what sorts of they take the headline ONS data [...] There's some core headline, maybe eight different facts and figures, sometimes nationally or sometimes regionally. So, what's the average spend for a household with a single dependent on food each month, cross referenced with all the stories on the application form, travel, average credit cards, that sort of classic stuff. Energy ONS has been traditionally used for a long time. However, in the last 12 months or so, ONS data is not keeping up with because they publish it typically once a year. I think they have increases or every six months now. But it's not keeping up with inflation.' (CR1)

When discussing alternative credit scoring mechanisms based on non-financial behaviours, such as internet browser or 'cookie' data (see Berg *et al.*, 2018), industry actors expressed scepticism. This was not due to doubts about the predictive power of these models but because there was no perceived need to shift away from current industry practices. Products like those from *Credolab* (<https://www.credolab.com/>), which provide lenders with behavioural data from devices and browsers - such as how often applicants input their income, take selfies or download finance apps - are already available. *Credolab's* approach integrates this data seamlessly into the customer journey, with consent obtained through standard terms and conditions, limiting the applicant's ability to opt-out. *Credolab* frames their use of behavioural data from smartphones and web metadata as a means to enhance financial inclusion and address issues with thin credit files (Credolab, 2023). While such data could potentially help prevent fraud - by identifying suspicious application speeds or detecting devices linked to previous fraudulent activities - these systems are not yet mainstream. Businesses commonly use 'cookie' data for targeted marketing, understanding the customer journey through social media ads and review sites, but integrating this data into credit scoring is less common. For example, one unsecured loan lender mentioned that while online behaviour might indicate desperation or caution in applicants, they had not incorporated this data into their credit models. If lenders are sceptical about the potential of Open Banking, they are even less likely to consider cookie data for credit scoring.

'I think it might be good at a higher level, but I just don't see lenders are going to be relying on that. When you get to the actual lending decision, the history of what they look at, like your profession, what your job, your career, your credit history, they work pretty well. In the mortgage industry, defaults on mortgages are actually really, really low, and so they're kind of, whilst people might want to say we need to disrupt it, actually, they do know what they're doing a lot of these lenders and building societies, in terms of actually the retail side. They know what good customers look like. Obviously, they're constantly trying to expand their criteria to help more people, but I don't think things like that would have a, yes, I'm not sure they'd have a huge impact because you're going to want to find the financial information anyway.' (CR3)

'There's nothing that is commonly used. And in fact, I don't know if any of our clients, they use any sort of data like that in our decision engines, because we help design and build them by the way to maintain them. We typically know every corner of the Decision Engine flowing through our data processing platforms. That sort of data that's consumer facing, by the way, so our clients manage the consumer facing application from the website, the app we're going through so I'm not aware of any of our clients using any of that to any sort of data.' (CR1)

Lastly, as an aside, and perhaps counterintuitively, one consultant explained how data indicating that customers have read the mortgage's terms and conditions is an applicant attribute associated with default that lenders cannot use in their models.

'And you, me, everybody. It's a hundred pages. You click to the bottom when you say, I accept. And off you go. Okay, a very small percentage. And you can see by how people use the process. A small percentage of people will read the terms and conditions, might be one [percent] something like that. Those people are massively high risk of default. [...] I think it's because that one big [indicator of] 'If I try and get out of this. But what's gonna happen?' So again, massively useful, I can't use it, because I can't say to the client 'Well, I would have lent to you, but you read the terms and conditions, so I'm not going to.' (Lender 4, consultant)

Conclusion

Building on the technological advancements in credit scoring over the past thirty years, the data available for making credit risk decisions continues to expand significantly. Lenders now incorporate non-financial service payment histories into their assessments and are increasingly seeking direct connections to essential data sources to expedite and simplify income verification. Open Banking technology is seen as having the potential to greatly enhance credit risk decisions, notably for those with thin credit files, and there is considerable activity in this area.

However, mortgage lending has been slower to adopt Open Banking, with some resistance due to concerns about transaction categorisation accuracy. There are also less regulated technologies that can digitally extract and analyse transactional data from bank statements in similar ways. While Open Banking applications may offer benefits such as marketing insights and early warnings of struggling borrowers, their full impact on mortgage lending remains uncertain.

What is clear is that the volume of digital data points available is increasing, which could potentially render traditional methods, like paper or PDF wage slips, obsolete through Open Banking or wage slip APIs. New banking transaction data provides deeper insights into personal lives and financial behaviours. Although Open Banking is gaining momentum and has proven successful in other lending markets, the future of what data will be essential for lending is still uncertain as the market transitions.

Chapter 5: Winners and losers of new technology in mortgage lending

Introduction

Being refused a mortgage is not always a negative outcome for applicants. The global financial crisis was partly driven by lenders who neglected to consider whether applicants could afford their loans, alongside aggressive marketing of mortgages, refinancing and additional borrowing (Ford and Wallace, 2009). Industry self-regulation failed to prevent these excesses, leading to the introduction of stricter regulations in 2014, which prioritised affordability assessments. These new rules, coupled with rising house prices, have made it more challenging for marginal buyers - particularly those with lower incomes, in routine jobs or self-employed - to enter homeownership. As a result, newer cohorts of homeowners are increasingly from professional occupations (Wallace et al., 2018). Stricter assessments of borrowers' ability to pay have enhanced the sustainability of homeownership, reducing the impact of the COVID-19 pandemic and other financial shocks on individuals and the housing market. It is also well documented that unsustainable homeownership can have lasting negative effects on mental and physical health, as well as disrupt social and family networks (Pevalin, 2009; Nettleton *et al.*, 1999). While homeownership offers benefits like wealth accumulation and security and can play a role in addressing wealth inequality (Christophers, 2017), it can also be harmful if not sustainable. Therefore, in some cases, being refused a mortgage might be the best outcome for certain applicants.

'Being declined isn't necessarily worse off? [...] It might be the best thing that's happened to you.'
(Lender 3, Consultant)

Nonetheless, in the USA, there have been concerns that the post-financial crisis thresholds for entry to homeownership – access to mortgage finance - are set too high (Goodman *et al.*, 2016) or that negative or conditional lending decisions fall more heavily on some groups rather than others (Wyly et al., 2009). Children of homeowners are more likely to become homeowners, aided as they are in significant ways by intergenerational wealth transfers, dubbed the 'Bank of Mum and Dad' (Christophers, 2017; Cook and Overton, 2023). What borrowers are supported by new credit risk assessment decision making tools therefore matters to who has access to homeownership and who does not. This chapter considers the project findings that relate to how new technologies influence what groups pass or fail mortgage lending tests.

Open Banking data

Beyond automating the administrative tasks of lending and using technology to detect potential fraudulent applications, a key innovation has been the rise of Open Banking technologies. As previously discussed, this technology analyses transactional details, as well as aggregated balances and flows from current account data, regardless of whether individuals have a positive or negative credit history. If Open Banking technologies are more broadly adopted, as is starting to happen in other lending sectors and with risk profiling tools in rental housing, there will be both beneficiaries and those who may lose out.

Open Banking is neither inherently better nor worse than traditional credit models based on payment history. While concerns about privacy intrusions and the social implications of Open Banking exist, the main difference lies in the type of information it provides for risk assessment. Open Banking reveals how individuals manage their finances, offering additional insights for credit risk managers even if the individuals do not fall into arrears or default. As discussed in the previous chapter, some companies position Open Banking technology as a solution for those with thin credit files, such as young people, migrants and even those who do not use credit but manage their finances well and are invisible to traditional financial services. These individuals are often directed to specialist lenders, despite potentially being creditworthy. One lender shared a story of an overseas banker who struggled to obtain credit in the UK, which, although possibly anecdotal, highlights the exclusion of those without substantial credit histories. In this context, Open Banking technology can be inclusive, potentially expanding access to homeownership compared to existing models, as well as to private renting, as noted in Report 3 and below.

'The head of technology for a big bank came and joined from America. It took him three years to get a credit card from that bank because he didn't have a credit footprint in this country; three years. He is a main board director of the bank [...]. He himself laughs to this day about it, that the only way he could access his own bank's credit card was with an improved credit score and it took him three years.' (CR4)

'I think it will also benefit, potentially, the people who haven't relied on credit. So, if you look at some subprime [borrowers] that are potentially in social housing do manage their finances well, but they do it without having any credit. At the moment they are disadvantaged because they're essentially doing a great job. They're not borrowing any money. They are making means and surviving on what they've got, but they're unfortunate to be in social housing. Now, if they want to step and go into a private rental, today they might struggle because they don't have any credit. So, it's those marginal cases where I think it will hugely benefit people. The other ones, which, in a broad demographic at the moment and are disadvantaged, are even professionals who are new to the country. Doctors for example who have lived and trained abroad and come to the UK to work but don't have any footprint in the UK and don't have any credit. They will struggle and yet they've got good jobs. They're in a professional career. They're paying and yet struggle to get a rental for that reason. So, I think there are pockets of outliers that hugely benefit from it.' (CR6)

Conversely, there will be losers for which Open Banking proves to be exclusionary. These could be people with sufficient salaries for affordability, who pay all bills on time but on closer inspection have a lot of debt, are always in an overdraft situation and do not manage their finances well. If borrowers sought mortgage finance from their current account providers, then this data was already used in decision-making, but for many people going to alternative mortgage lenders, where once they would have benefits in the future, they may find themselves in a less advantageous position, penalised, refused a loan or offered one on poorer terms.

'Yes, so from a credit report perspective, you could have someone who essentially has a lot of borrowing. So, they could have four or five credit cards and a loan, and their income could pass any income tests and when you look at the credit profile, they're paying all their bills on time, but what Open Banking will allow you to see is if they're using credit to pay for credit. So, if they were potentially taking out loans to pay off credit cards and then doing balance transfers to pay the other one and the whole time, they're continuously in their overdraft, Open Banking will allow a lender to spot that, which ordinarily might have passed any traditional credit check. They're on the voters' roll. They've passed the income ratios, and they pay their bills on time, but the reality is they're one month away from blowing up because they don't have any savings and they're using one borrower to pay another borrower. When they go bad, they go bad because they've got a trail of debt. So, they're the ones where I think potentially the lender benefits more than them.' (CR6)

'You should have got a better offer from your existing, or a more accurate offer from your existing current account provider, if you were in the space to get a good offer candidly, if you weren't, if your current account wasn't in great shape, but you still needed a mortgage. So, if you were a little bit marginalised, though you know in financial difficulties, you probably didn't want to go to your existing current account provider, and you would hope that the mortgage provider isn't looking at this stuff, cause you desperately need a place. But your current account isn't saying you can afford it.' (CR7)

Inclusionary practice

One industry insider considered that tighter mortgage market regulation has made the mortgage market more exclusionary. Automated systems may reduce costs for the mainstream, but marginal households cannot be accommodated in these systems, and many are rerouted towards more expensive specialist firms.

'There've been a number of policy failures in banking over the course correct [reregulation after the global financial crisis], but probably the biggest and most harmful to customers...Policy failure as being ringfenced banking just basically meant that if you are a consumer who is not complicated, and when I say complicated, I can mean your couple on one of you is foreign, can mean you haven't been in your current job for the year. It can mean you share your name with a prominent terrorist. It could mean that your parents are Iranian. It could mean your parents are a politician, let alone it can mean you had a bit of a difficult credit history in the past. If any of those things flag up, then it is economically impossible for a major bank to give you a mortgage. So, you're given to a small specialist provider of what a specialist provider does, and there are thousands of them, is they deal with that and get themselves comfortable that will lend to you and that cost them a lot of money, and they charge the consumer a lot of money for doing [so]. That's kind of what happens.' (Lender 4, Consultant)

Participants noted how lenders can, however, make a staged decision process and might try different data sources to give people chances to obtain credit, but this sort of process is more costly and complicated than getting all the information up-front and offering simple accept or decline decisions. Such staged systems that provide a pathway to Open Banking may give lenders a robust defence of their decision-making but would make these more complex systems expensive and obtaining additional data may be problematic, if borrowers are reluctant to engage with Open Banking, for example.

'Therein lies the challenge in that the way some people and certain companies have built that process is if they don't get consent, then they're declined from that application. This comes down to how you bespoke the application process for the person that's in front of you. So, yes, the optimal dataset is Open Banking, but should companies lead with that as the request or should they try and get to a good answer on the data that they've got and only if it's a fringe case or it's looking like a decline should they then go on to request further information? So, I think there's potential for lenders to do those custom journeys for the clients, but that's costly.' (CR6)

'All of this is even more important in the current regulatory environment, when we look at Consumer Duty, which is essentially based around outcomes, and making sure that customers have suitable products, so not just products that they can afford, but products that they'll continue to be able to afford.' (CR2)

This situation often arises manually, particularly for self-employed borrowers who face extra scrutiny to prove their income. During the pandemic, some self-employed individuals had no income at that specific time, making it even more challenging. Despite having audited accounts and years of income records, one such borrower interviewed had to provide additional

documentation and meet stricter criteria. In an automated model, an application like this would likely be immediately rejected, sent to a manual underwriter, or in more advanced systems, directed through a staged process that incorporates additional data from sources like Open Banking or HMRC. This additional data helps provide further assurance of the borrower's financial history, circumstances or income.

A credit risk software company mentioned they aimed to help place loans for individuals in these ambiguous cases. While high-income contract workers usually posed no significant issues, lenders were more cautious with those in less secure and lower-paid jobs. This firm worked with lenders to broaden their criteria for discretionary inclusion and supplied data to support the applications of such borrowers.

'What we've got is we've got quite a lot of, so the lenders will publish their criteria, but a lot of the lenders also have a grey area, which they don't publish, where they will push the boundaries, the criteria. What we do is we've got all those grey areas, where we know we can push the limits with the underwriters to help people get through.' (CR2)

Some interviewees maintained that there should be a greater focus on widening participation in lending, highlighting that discretion and opportunities often granted to more affluent borrowers are not extended to lower-income borrowers. For example, marginal cases involving high-value lending are frequently passed on for manual underwriting, while similar cases involving low-income borrowers are often rejected without a deeper appraisal of the individual's circumstances. Research shows a social gradient in mortgage arrears, with borrowers in the lowest income decile being twice as likely to have arrears compared to those in the highest income decile (Wallace et al., 2018). But interviewees suggested that this is an area where additional data or human intervention could be valuable in expanding access to homeownership.

One firm noted that their use of Open Banking technology could significantly improve fairness in lending decisions, addressing the issue of thin credit files that prevent many from becoming homeowners. Rather than using technology solely to price or minimise risk through exclusion, they argued that technology can also be used to better match borrowers with suitable products, using holistic data to ultimately support more favourable outcomes.

'There should be specialists in almost poverty lending. Really, it's marginal cases is 'Okay, Where's the, what are we doing here? With the risk?' And actually, looking at a full financial statement as an individual, and saying, if we do it, if we do a consolidation of your finances and do that properly, [...] If we can help you with that consolidation, if we can use your mortgage to pay off a personal loan and a car loan or mug up student debt, or whatever it is. Does that change you? Financial position? [...] Can we look at; can we look at you as an individual thing? You have just come out of university, what's your career, prospects over the next 3, 4, 5 years? Okay, it just makes sense. Now, you're on the staff of something great, and we want to be on that journey with you. That's the human element for me.' (CR7)

'My takeout, whether it's right or wrong, from the few conversations that I have had [on Open Banking], is that there's more of an ethical thing going on here around giving a sort of democratisation of everyone should have equal rights to get a loan kind of thing. So obviously on normal scoring it seems relatively unfair, even actually if you're relatively wealthy, because you haven't got credit, you can't get a loan potentially. Then the other extreme, and that's where I think [firms name] really exciting, what I took out of it at least.' (CR2)

'I think for me we hear it a lot around placing the customer first. The thing with lending is the customer often isn't put first. The motivation of 'will they pay me back or not?' is the determining

factor. It is starting to edge that a little bit, but it needs to be more about, as you say, what is the right product and therefore what tools and technology do I need to find the right product as opposed to are they in or are they out? That does bring up a topic that we've moved away from as an industry which is risk-based pricing, which isn't good for anybody either. So, there is a real fine line because risk-based pricing just infers that people who need the help and are potentially struggling end up paying more for borrowing than the people who can afford it. So, there's that moral duty that sits behind it as well.' (CR6)

One emergency lender showcased the potential of algorithmic lending by integrating a benefits calculator into their online systems. Although it is unclear how this was organised - as benefit eligibility and entitlement may be hard to glean from income and expenditure data alone, without knowing the circumstances of family members, for example - this allowed the lender to identify significant amounts of unclaimed benefit entitlements among their working clients, highlighting a valuable opportunity that other lenders could explore.

'The best thing that we did last February was to put a benefits calculator in our loan journey. [...] We have identified - and I'm glad you're sitting down and breathing normally, £324 million pounds worth of unclaimed benefits since February of last year. Which is absolutely staggering. We were the first people in the UK to do that and it's now been widely adopted by other CDFI lenders...[...] In fact, the fact that we can identify benefits for them, we're going one step further, we're also trying to identify any grants or other things that they could be regionally entitled to' (CR4)

Bias and protected characteristics

As discussed in Chapter 3, although technology is currently in the spotlight for its capacity to replicate and amplify existing social inequalities (Eubanks, 2018; O'Neill, 2016) human decision-making is fallible too. This section looks at participants' views of how lenders consider these risks.

Credit risk technology firms occasionally acknowledged that they were aware of inadvertently introducing bias into new models, which were difficult to test without deploying them live. For example, using historical data as training data can replicate past inequalities, thereby introducing bias into the new systems. From this limited set of interviews, it was not clear that testing for bias was a common practice, and the appropriate steps to address any identified bias were also uncertain.

'So, I've done bits on like the pricing side. And if you think about how, you kind of work out the pricing of personal loans, etc., the headline rate is always really low for a personal, so that the best customers get the, and then the poorer customers get higher rates because of that's their credit risk. If you were to look at ethnicity, or gender, or anything through that, and there was a skew. I don't know what you would do. I don't know.' (Lender 3, Consultant)

Gender is a key protected characteristic, and several participants raised concerns about its role in car insurance that could have wider implications. While charging different premiums based on gender has been prohibited under equalities legislation, there are fears that other indicators of gender, or other protected characteristics, could still be present in ML models. This raises concerns about how the industry safeguards against such biases. One consultant recounted being asked to evaluate an ML-based credit risk model used by an overseas lender. The model had identified military service as a significant risk factor, noting that men were at a higher risk of default due to the country's conscription system. Participants acknowledged that these risks can arise when proxy indicators of protected characteristics inadvertently make their way into models.

'Lending default [is] correlated to gender but can't use it, same with postcode but can't use it in the models and if they're linked in, or the models they use in other ways by using a simple model that doesn't have gender in it. And then you never have the conversation. You never have the conversation that says with the other factors in the model, to what extent does this replicate gender? And, by the way, whether it's car insurance, or whether it's lending, gender is there. They are always there. A factor can be correlated, you know, or a factor can be causative, or a factor can provide marginal information. which means you can't pick it up with other factors. And the truth is and again, [...] gender does provide marginal information. I can build a better model if you allow me to use gender, but you can't. So, people don't, same with postcodes, postcodes are fantastic. And you can't use it. But first of all, you can't say, well, we would have lent to you, your neighbour defaulted last year. So, we're not going to, feels wrong, and the other thing is, of course, that the street, the user one, is massively highly correlated to your ethnic group, so just looking at it, it will, it will discriminate, based on that.' (Lender 4, consultant)

'So, the most noticeable bias that the credit risk or decisioning industry had was around car insurance whereby gender used to be a key rating factor of your insurance. It was well-known that a female driver had a lower premium because insurers would use gender as a rating factor. Now, now [gender] being a protected characteristic, insurers are no longer allowed to set your premium using gender. That was I guess the first of that kind of misuse of a protected characteristic. It's whether machine learning and AI bring that back into play and unconsciously create a bias. It doesn't necessarily have to be protected. There is a pattern whereby 44-year-old males living in Grimsby never pay their credit card, but machine learning and the model would therefore dictate that any other 44-year-old male living in Grimsby applying for credit would be a decline. So that's where the bias of machine learning and AI can come into play, and we need to be careful about how we protect against that.' (CR6)

As in the private rented sector, self-employed people regularly had difficulty proving their income and the security of their income - more understandable, perhaps, in a long-term high-value relationship such as mortgage lending compared to a one-year tenancy. One emergency short-term loan lender focussed on public sector workers as they were often low-paid but also were more likely to have secure employment. Some self-employed mortgage borrowers had been asked to provide a guarantor to obtain the loan. The guarantor was a work colleague who had to undertake the same affordability and financial scrutiny as the borrower. One firm was lending specifically in this area, using an automated platform to combine affordability and stress testing of applicants and their parents, using family assets like home equity and pensions to boost the applicant's affordability assessments. Their platform links immediately to the market and can provide rapid decisions in principle and a more inclusionary practice, but one that is conditional on family circumstances and, as in renting, spreads risk beyond the immediate borrower (or tenant).

'My guarantor is someone we work with in school... because they do understand you, they will just finally do it to assist you.' (MB2)

'Our really classic user-case is someone whose family might be the first generation to own a home, and actually they're seeing that their children might actually not be able to buy a home because they can't give them the support. It's those kinds of people who, they might have some assets, but they don't have cash just to hand out. They're not wealthy. They can give their families that stepping stone. That's one group of products. [...] We've built some technology that allows us to do all the affordability and all the assessment, and we work with all these partners to get people on to the property ladder.' (CR3)

The use of conditional terms, such as requiring guarantors, is less common in the mortgage market compared to the private rented sector, but some products aimed to draw in family

resources into affordability assessments. Typically, however, instead of having an alternative payees, mortgage borrowers who are deemed higher risk by assessments are more likely to be placed on higher-cost loans to manage that risk. As identified in the literature, credit risk assessments have shifted from merely mitigating risk to actively managing it, moving away from excluding borrowers towards ensuring profitability. Aalbers (2011) describes this approach as 'yellow lining,' where borrowers are included but under unfavourable terms, such as higher-cost or subprime loans. This is in contrast to 'redlining,' which involves the outright exclusion of people based on certain circumstances or geographic locations.

Borrowers interviewed were refused mortgage loans due to debt histories or found lenders who would advance loans but on higher interest rates and/or with a guarantor. Another interviewee described the difficulties in getting a mortgage due to her partner's poor credit history and her self-employment status. She noted that her partner had defaulted on credit cards about five years ago, impacting his credit score and thus their ability to get a joint mortgage. Subprime lending was initially welcomed to give people the chance to rehabilitate their credit files but has declined as tighter regulation supports affordability and stress testing of payments.

'In the second company, they did consider me, but then it had higher interest rates, and I also had to bring in a guarantor for me to qualify for the mortgage.' (MB2)

'I got a mortgage by myself without him because they wouldn't give him one, and that was all due to his credit, and his bad credit. [...] He's been working really hard to get his credit score up, and regularly checking all of the Experian and Equifax and all of the agencies in order to improve it.' (MB1)

One lender, reliant on sophisticated AI/ML models, highlighted his underwriters' different subjective takes on cases, exposing the biases of human decision-making, but emphasised how these judgements feed into algorithmic models. They use Open Banking data for emergency loans for public sector workers and found many people doing *Omaze* [<https://www.omaze.com/>] type raffles and giveaways. They took a view to label this gambling, and they exclude all applications where gambling is present.

'Clearly, there is bias and noise in all decisions that are being made. If I have Tracey who's 55 decisioning a young male and then, I have Tom who's 30 decisioning the same male, they'll come up with two different answers. [...] Alison, would you rather lend money to someone to buy a clarinet or a PlayStation 5? You've already made your mind up. Q: *It's just everybody's right to do either, isn't it, really?* Well, there is but already the bias is I'd rather lend to someone that was playing a clarinet. It's just complete rubbish. I think it works both ways. I think age is a big factor in lending, definitely. The time of day is a big factor. The time that someone applies for credit.' (CR4)

Credit scoring is a form of financial and social discipline, and now Open Banking and related technologies could become additional tools of financial control. Currently, awareness of this is limited, but how we present our current account data may soon be as crucial to financial opportunities as managing our credit scores. There is already some resistance to a cashless society and concerns about banks knowing the most private details of our lives through our digital spending habits (Scott, 2022; Donohue *et al.*, 2020). However, even without abandoning the convenience of electronic payments, people will need to be more mindful of how they present their transactional data to a financial services industry that increasingly sees more about the population. Just as financial education apps provide access to credit scores and consumer advocacy offers guidance on managing credit history, managing the patterns in our current accounts is likely to become equally important. One key criticism of algorithmically determined decision-making is its lack of transparency. This opaqueness is crucial because if

people do not understand how firms make decisions, they cannot effectively assert their rights to fair treatment under consumer protection and equalities legislation (Selbst and Powles, 2017).

'Yes, so one of the key things around where you use anything like machine learning or AI is to have that explainability, so to have that audit trail of why you made that decision. That's as far as it's got at the moment. No one I believe has been prosecuted or fined for misuse of AI or machine learning at the moment. I'm not sure how the regulator will police it other than your obvious biases.' (CR6)

Mortgage futures

One lender believed that future mortgage lending might be influenced by ongoing reliance on technology and automation, combined with a market structure dominated by a few large players. These trends could lead to a more streamlined and efficient mortgage process, but might also reduce innovation and flexibility due to increased regulation and market consolidation. However, a building society lender noted that despite the rise of technology setting fixed parameters for loans, there will still be opportunities for smaller lenders. This is because not all individuals fit neatly into these standardised models.

'I can only think that there will always be a manual review or something, because as much as tech is brilliant at the moment. And there's more and more data available to us as a lender. I think it's; you'll always get a customer that's just falling outside of that.' (Lender 5)

Despite some scepticism from long-standing industry insiders, Open Banking is seen as a significant part of the future of mortgage lending, along with other banking innovations. One such innovation is Open Finance, which extends the concept of consumer data ownership and sharing beyond current accounts to other financial products, such as pensions. This would provide additional data on wealth, debts and income, which is relevant for lending, especially into retirement. Other innovations include the digitisation of the home buying process. Some digital mortgage brokers are preparing to offer online conveyancing services, while others are integrating with property searches.

'So, my initial view on this is we have to, the immediate next step is to get to Open Finance, and making sure that everybody in the marketplace is playing ball with that data, which at the moment is not the case. I was reading, I think in an Open Banking expo paper, a few days ago that Brazil actually is one country that is already there with Open Finance.' (CR6)

'I think some of the work around conveyancing, so another completely archaic process is around how all the local authorities have, in effect, physical copies of all the Land Registry searches in their offices, which is why conveyancing takes six weeks. We're actually working on a product, which is in beta at the moment, whereby our customers can come on to our website and they log in, and once they've gotten to a stage where they might start looking for properties, they can type in the address of the property and we will, in effect, do the conveyancing search that a solicitor would do for them.' (CR3)

Conclusion

Algorithmic changes in data resources and credit risk decision-making represent a significant technological shift in the mortgage market, likely to result in both advantages and disadvantages for different individuals. A broader range of financial and behavioural data sources can lead to more accurate credit risk assessments, but this increased accuracy might benefit some people

while disadvantaging others, potentially impacting perceived fairness. Systems could be designed to be more inclusive by requesting additional data that reflects good financial management without automatically rejecting applications in mainstream lending. Manual underwriting might still be used for marginal low-income cases as well as for applications from more affluent households. While there is awareness of the risk that algorithms could adopt or reflect bias, there is limited insight into how to manage and mitigate this risk. Despite some scepticism in the mortgage lending industry, there is a trend towards greater use of digital data and automation. With Open Banking gaining momentum, if it becomes significant, there will be a need for financial education to include guidance on how consumers can manage and present their transaction data effectively to the financial services industry, similar to how people are currently educated about managing their credit scores.

Chapter 6: Discussion and conclusion

In the rapidly evolving landscape of financial services, credit data is no longer confined to traditional banking transactions. The scope of data resources is expanding beyond banks to include utilities, telecoms and other non-financial sectors. This shift is driven by advances in digital memory and analytical capabilities, allowing institutions to handle and analyse vast amounts of data with unprecedented efficiency.

The automation of credit assessment processes is at the heart of this transformation. Traditionally, assessing creditworthiness involved manual review of numerous factors, including detailed financial histories and personal circumstances. Today, increased automation is streamlining these processes, particularly for high-volume lending scenarios. Automated systems can handle large numbers of standard cases efficiently, but they still encounter challenges when dealing with more complex or non-standard applications. For such cases, manual intervention remains necessary unless lenders invest in sophisticated, multi-stage software. While these advanced systems can be costly, ongoing technological advancements may soon make them more affordable and widely accessible.

One of the most significant innovations in credit assessment is Open Banking. This technology leverages detailed banking transaction data to evaluate a borrower's financial health, providing a comprehensive view that goes beyond traditional credit scores. Open Banking has the potential to revolutionise the industry by addressing common issues such as thin credit files and incomplete financial profiles. By analysing transaction data, lenders can gain deeper insights into a borrower's income, expenditures, and overall affordability.

However, Open Banking is not without its challenges. One notable issue is the accuracy of categorising spending as either essential or discretionary. This distinction can significantly impact credit assessments, especially at the margins where financial decisions are most sensitive. Open Banking can reveal patterns that traditional credit scoring might miss, such as high-income individuals who poorly manage their finances or frequently use overdrafts. Despite these insights, some lenders still rely on alternative methods, such as analysing PDF bank statements, to automate data review and enhance the efficiency of manual analysis.

The adoption of Open Banking and similar technologies has sparked a range of reactions within the lending industry. Agile, innovative firms often champion the potential of these new tools, envisioning a future where data-driven insights lead to more precise and equitable credit assessments. In contrast, long-standing industry insiders may exhibit scepticism, preferring to stick with established systems that, while less innovative, are tried and tested. This hesitation mirrors the initial resistance to credit scoring automation, which eventually became a standard practice in the industry.

As the industry adapts to these new technologies, it remains unclear how they will be integrated into existing systems. Lenders might access these tools directly through FCA authorised third parties, or they could be incorporated into broader data aggregation systems managed by credit reference agencies. This uncertainty reflects a broader trend in the financial sector: the shift towards greater automation and data integration is inevitable, but the exact pathways and implications are still unfolding.

Consumer awareness is another critical aspect of this evolution. While many borrowers understand the importance of their credit scores, they may be less aware of how their

transaction data is used by financial institutions. Open Banking has the potential to create both winners and losers, highlighting the need for improved public education and financial guidance. The recent FCA (2023) *Credit Information Market Study* emphasises the need to increase consumer awareness not only of traditional credit information but also of emerging data sources like Open Banking.

The ethical and fairness considerations surrounding automated credit assessments are also worth noting. Training data used to develop credit scoring models can inadvertently perpetuate existing biases, including those related to protected characteristics. Although lenders are aware of legal requirements to avoid discriminatory practices, the effectiveness of their efforts to address these biases can vary. Some lenders employ staged models that increase data requests for marginal cases, but inconsistencies in manual underwriting and reliance on specialist lenders can complicate efforts to ensure fair outcomes. This reflects a broader concern about how data-driven decisions impact fairness and accessibility, particularly for those in more marginal financial situations.

The broader implications of these technologies extend beyond individual lending decisions. Homeownership plays a significant role in wealth accumulation and life opportunities, and fair access to credit is crucial for reducing wealth inequalities. As the financial sector moves towards greater automation and data-driven decision-making, examining how these changes impact fairness and accessibility will be essential.

In conclusion, the landscape of credit assessment is undergoing a profound transformation driven by advancements in technology and data integration. Automation and Open Banking are reshaping how lenders evaluate creditworthiness, offering new opportunities and challenges. As these technologies become more embedded in the industry, their impact on consumers, lenders, and the broader financial system will continue to evolve. Ensuring that these changes promote fairness, transparency, and accessibility will be crucial for achieving positive outcomes in the ever-changing world of credit and lending.

Recommendations

Across the Code Encounters project, we identified universal themes that need to be addressed as well as sector specific that require attention, including the following that are relevant to Government, those responsible for financial education, risk profiling technology firms, trade bodies, lenders, landlords and agents.

- 1. To make visible how data and algorithms have been used in each decision** Provide greater transparency in the way data is gathered from and about borrowers and offer clarity about how these data will be used.
- 2. To establish agreed guidelines on the appropriate use of algorithms for stakeholders within the sector and tenures** Provide guidance to lenders on what new data and automation can do and how algorithms and new data resources may be deployed.
- 3. To produce guidance on the use of data and algorithms for borrowers** Increase public awareness of how they must manage their digital profiles, including banking

transaction data, much in the same way as the importance of managing credit scores has permeated financial education and public consciousness.

4. **To retain human oversight in decision making** Not all people fit algorithmic models so human oversight should be maintained to ensure fairness.
5. **To ensure the explainability of decision making** Organisations must be able to fully articulate how a decision was reached, including the data used, where algorithms were involved and the human oversight of the outcome.
6. **To ensure the retention of flexibility and individually tailored decision-making** We would suggest having a system in place in which the inputs into algorithmic processing can also be adapted to enable flexibility and to ensure that both input and outcomes remain flexible and adaptable to the individual being assessed.

These recommendations are discussed in more detail in our *Overarching summary report 1*.

References

- Aalbers, M.B. (2011) *Place, Exclusion and Mortgage Markets*. John Wiley & Sons.
- Adkins, L. (2018). *The Time of Money*. Stanford: Stanford University Press.
- Aggarwal, N. (2021). The Norms of Algorithmic Credit Scoring. *The Cambridge Law Journal*, 80 (1), pp.42–73.
- Amoore, L. (2020) *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Duke University Press.
- Arya, S., Eckel, C. and Wichman, C. (2013). Anatomy of the credit score. *Journal of Economic Behavior & Organization*, 95, pp.175–185.
- Berg, T., Burg, V., Gombović, A., et al. (2020) On the rise of FinTechs: Credit scoring using digital footprints. *The Review of Financial Studies* 33(7) 2845–2897.
- BoE and FCA. (2019). Machine learning in UK financial services. Bank of England and FCA . [Online]. Available at: <https://www.fca.org.uk/publication/research/research-note-on-machine-learning-in-uk-financial-services.pdf>. [Accessed 15 February 2022].
- Bono, T., Croxson, K. and Giles, A. (2021) Algorithmic fairness in credit scoring. *Oxford Review of Economic Policy* 37(3). Oxford Academic: 585–617.
- Brick Underground (2015) *Redlining: How one racist, depression-era policy still shapes New York real estate*. http://www.brickunderground.com/blog/2015/10/history_of_redlining Accessed 16 August 2024.
- Bücker, M., Szepannek, G., Gosiewska, A. and Biecek, P. (2022). Transparency, auditability, and explainability of machine learning models in credit scoring. *The Journal of the Operational Research Society*, 73 (1), pp.70–90.
- CDEI. (2020). Review into bias in algorithmic decision-making. [Online]. Available at: <https://www.gov.uk/government/publications/cdei-publishes-review-into-bias-in-algorithmic-decision-making/main-report-cdei-review-into-bias-in-algorithmic-decision-making> [Accessed 20 April 2022].
- Cevolini, A. and Esposito, E. (2020). From pool to profile: Social consequences of algorithmic prediction in insurance. *Big Data & Society*, 7 (2), pp.1–11.
- Cheeseman, S. (2024) Top UK Challenger Banks | 2024 Beauhurst Blog. <https://www.beauhurst.com/blog/uk-challenger-banks/> Accessed 22 August 2024.
- Citron, D.K. and Pasquale, F. (2014). The scored society: Due process for automated predictions. *Washington Law Review* , 89, p.1.
- Cook, J. and Overton, L. (2023). Intergenerational Assistance with Home Ownership: Understanding the Relational Development of Financialized Subjectivities. *Housing, Theory and Society*, 41(2), 216–233. <https://doi.org/10.1080/14036096.2023.2293188>

Credolab (2023) *6 Ways To Utilise credolab's Behavioural Solutions to Drive Financial Inclusion* <https://www.credolab.com/blog/6-ways-to-utilise-credolabs-behavioural-solutions-to-drive-financial-inclusion> Accessed 22 August 2024.

Christophers, B. (2017) Intergenerational Inequality? Labour, Capital, and Housing Through the Ages. *Antipode*. DOI: 10.1111/anti.12339.

Ciocănel, A., Wallace, A., Beer, D., Cussens, J., and Burrows, R. (2024). Open Banking and data reassurance: the case of tenant referencing in the UK. *Information, Communication & Society*, 27 (10) <https://doi.org/10.1080/1369118X.2024.2310481>

Dastile, X., Celik, T. and Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, p.106263

Deville, J. (2020). Futures of credit risk assessment in the UK. [Online]. registry-trust.org.uk. Available at: https://www.registry-trust.org.uk/documents/46/Futures_of_Credit_Risk_Assessment_-_March_2020.pdf [Accessed 1 February 2022].

Deville, J. and Van der Velden, L. (2015). Seeing the invisible algorithm: The practical politics of tracking the credit trackers. In: *Algorithmic Life*. London: Routledge. pp.101–120.

W. Donohue, Z. Afridi, K. Sokolyuk, T. Bedwell, E. R. York and A. A. Salman, (2020) "Cashless Society: Managing Privacy and Security in the Technological Age," *Systems and Information Engineering Design Symposium (SIEDS)*, Charlottesville, VA, USA,, pp. 1-6, doi: 10.1109/SIEDS49339.2020.9106653.

Equifax (2024) *The increasing importance of accurate affordability checks.* https://www.equifax.co.uk/about-equifax/blog/all-news/en_gb/-/story/the-increasing-importance-of-accurate-affordability-checks?utm_source=eloqua&utm_medium=email&utm_campaign=oth_awareness_newsletter_010824&utm_content=aug24news Accessed 22 August 2024.

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Publishing Group.

Experian (2018) *Making the Invisible Visible: Exploring the power of new data sources*, London: Experian UK

Experian (n.d.) How categorisation and aggregation bring Open Banking to life <https://www.experian.co.uk/blogs/latest-thinking/automated-credit-decisions/how-categorisation-and-aggregation-bring-open-banking-to-life/> Accessed 28 June 2024.

FCA (2012) *Mortgage Market Review Data Pack: Supplement to PS12/16*. London, Financial Conduct Authority.

FCA (2022) *Credit Information Market Study Interim Report and Discussion Paper Market Study MS19/1.2* London, Financial Conduct Authority. <https://www.fca.org.uk/publication/market-studies/ms-19-1-2.pdf> Accessed 22 August 2024.

FCA (2022a) A new Consumer Duty Feedback to CP21/36 and final rules. London, Financial Conduct Authority. <https://www.fca.org.uk/publication/policy/ps22-9.pdf> Accessed 22 August 2024.

FCA (2023) *Credit Information Market Study Final Report: Market Study MS19/1.3* London, Financial Conduct Authority. <https://www.fca.org.uk/publication/market-studies/ms-19-1-3.pdf> Accessed 22 August 2024.

Foohy, P. and Greene, S. S. (2021). Credit Scoring Duality. *Law and Contemporary Problems*, Forthcoming 2022. [Online]. Available at: doi:10.2139/ssrn.3992749 [Accessed 13 April 2022].

Ford, J. and Wallace, A. (2009) *Uncharted Territory?: Managing Mortgage Arrears and Possessions*. London, Shelter.

Fourcade, M. and Healy, K. (2013). Classification situations: Life-chances in the neoliberal era. *Accounting, Organizations and Society*, 38 (8), pp.559–572.

Fourcade, M. and Healy, K. (2017) Seeing like a market. *Socio-Economic Review* 15(1). Oxford Academic: 9–29.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T. and Waither, A. (2022) Predictably Unequal? The Effects of Machine Learning on Credit Markets. *The Journal of Finance* LXVII(1): 5–47.

Henry, N., & Morris, J. (2018). Scaling Up Affordable Lending: Inclusive Credit Scoring. *Responsible Finance*. <https://www.european-microfinance.org/publication/scaling-affordable-lending-inclusive-credit-scoring> Accessed 22 August 2024.

Hohnen, P., Ulfstjerne, M. A., & Krabbe, M. (2021). Assessing creditworthiness in the age of big data: a comparative study of credit score systems in Denmark and the US. *Journal of Extreme Anthropology*, 5(1), 29-55.

Hollinrake, K. (2024) The role of Smart Data in improving the home buying and selling process in England (Letter to Clive Betts MP) <https://committees.parliament.uk/publications/44208/documents/219945/default/> Accessed 22 August 2024.

Hurley, M. and Adebayo, J. (2016) 'Credit Scoring in the Era of Big Data.' *Yale Journal of Law and Technology*, 18, 2016, pp. 148-216. HeinOnline.

IMLA (2022) Mortgage market to slow in 2023 and 2024 but borrowers remain resilient in the face of global upheavals, IMLA report predicts. 21 December 2022 <http://www.imla.org.uk/news/post.php?s=2022-12-21-mortgage-market-to-slow-in-2023-and-2024-but-borrowers-remain-resilient-in-the-face-of-global-upheavals-imla-report-predicts> Accessed 22 August 2024.

Johnson, M. (2021) Challenger banks are dead, long live challenger banks. *Journal of Digital Banking* 5(4). Henry Stewart Publications: 329.

Kassab, M. and Laplante, P.A. (2022) Open Banking: What it is, where it's at, and where it's going. *Computer* 55(1). Institute of Electrical and Electronics Engineers (IEEE): 53–63.

Kear, M. (2017) Playing the Credit Score Game: Algorithms, 'positive' data and the personification of financial objects. *Economy and Society*, 46(3-4), pp.346–368.

Keasey, K. and Veronesi, G. (2012) The Significance and Implications of being a Sub-Prime Homeowner in the UK. *Environment and Planning A* 44 pp.1502-1522.

Kim, S.D., Andreeva, G. and Rovatsos, M. (2023) The double-edged sword of big data and information technology for the disadvantaged: A cautionary tale from Open Banking . *arXiv [cs.LG]*. Available at: https://www.crc.business-school.ed.ac.uk/sites/crc/files/2024-06/the-double-edged-sword-of-big-data-and-information-technology-for-the-disadvantaged_0.pdf. Accessed 20 August 2024.

Kitchin, R. (2017) Thinking critically about and researching algorithms. *Information, Communication and Society* 20(1)14–29.

Langley, P. (2009). Debt, Discipline, and Government: Foreclosure and Forbearance in the Subprime Mortgage Crisis. *Environment & Planning A*, 41 (6), pp.1404–1419.

Langley, P. (2014). Equipping entrepreneurs: consuming credit and credit scores. *Consumption, Markets & Culture*, 17 (5), pp.448–467.

Lauer, J. (2017). *Creditworthy: A History of Consumer Surveillance and Financial Identity in America*. Columbia University Press.

Leyshon, A. and Thrift, N. (1999). Lists come alive: electronic systems of knowledge and the rise of credit-scoring in retail banking. *Economy and Society*, 28 (3), pp.434–466.

Lyon, D. (ed.) (2002) *Surveillance as Social Sorting*. Taylor & Francis Group.

Marron, D. (2009) *Consumer Credit in the United States: A Sociological Perspective from the 19th Century to the Present*. New York : Palgrave Macmillan.

Munro, M., Ford, J., Leishman, C. and Kofi Karley, N. (2005) *Lending to higher risk borrowers Sub-prime credit and sustainable home ownership*. York, Joseph Rowntree Foundation.

Nettleton, S., England, J., Burrows, R. (1999) *Losing the Family Home Understanding the Social Consequences of Mortgage Repossession*. York: York: York Publishing Services Ltd.

Onay, C. and Öztürk, E. (2018). A review of credit scoring research in the age of Big Data. *Journal of Financial Regulation and Compliance*, 26 (3), pp.382–405.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Books.

OPA (2024) 'Justice Department Secures Agreement with Patriot Bank to Resolve Lending Discrimination Claims' Press Release. 17 January 2024. Office of Public Affairs US Government.

<https://www.justice.gov/opa/pr/justice-department-secures-agreement-patriot-bank-resolve-lendi>

[ng-discrimination-claims#:~:text=Since%202021%2C%20the%20department%20has.lending%20discrimination%20across%20the%20country. Accessed 22 August 2024.](#)

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge and London: Harvard University Press.

Pevalin, D.J. (2009) Housing repossessions, evictions and common mental illness in the UK: results from a household panel study. *Journal of Epidemiology and Community Health* 63(11): 949–951.

Poon, M.A. (2008) From New Deal Institutions to Capital Markets: Commercial Consumer Risk Scores and the Making of Subprime Mortgage Finance. [https://papers.ssrn.com](https://papers.ssrn.com/papers) › [papers](https://papers.ssrn.com/papers)<https://papers.ssrn.com/abstract=1458545> (accessed 13 January 2022).

Raffa, M., 2024. Make-Do-With Listening: Competence, Distinction, and Resignation on Music Streaming Platforms. *Social Media+ Society*, 10(1), p.20563051231224272.

Rae, A. (2015), 'The illusion of transparency: the geography of mortgage lending in Great Britain', *Journal of European Real Estate Research*, Vol. 8 No. 2, pp. 172-195.
<https://doi.org/10.1108/JERER-08-2014-0030>

Reid, C.K., Bocian, D., Li, W. and Quercia, R. (2017) Revisiting the subprime crisis: The dual mortgage market and mortgage defaults by race and ethnicity. *Journal of Urban Affairs* 39(4). Informa UK Limited: 469–487.

Reynolds, F. (2021). *Open Banking. A consumer perspective*. Barclays.
[https://doi.org/10.1016/s0378-4274\(21\)00410-0](https://doi.org/10.1016/s0378-4274(21)00410-0)

Rona-Tas, A. (2017). The Off-Label Use of Consumer Credit Ratings. *Historical Social Research / Historische Sozialforschung*, 42(1 (159)), 52–76.

Salad Money (2023) *Credit where Credit is Due*. London, Salad Money.
<https://www.saladmoney.co.uk/about/reportdownload?reportID=2> Accessed 25 September 2024.

Schmeckpeper, K., Roberts, S., Ouellet, M., Malencia, M., Jain, D., Gosrich, W. and Bromberg, V. (2021). Algorithm transparency through the Fair Credit Reporting Act (FCRA). *Journal of Science Policy & Governance*, 18 (04). [Online].

Scott, B. (2023) *Cloudmoney: Cash, Cards, Crypto and the War for our Wallets*. Spain: Penguin Random House Grupo Editorial.

Selbst, A.D. & Powles, J. (2017) Meaningful Information and the Right to Explanation. *International Data Privacy Law* 7(4),pp. 233-242

Sparkes, M. (2024). Housing, debt, and credit score classification situations: what the unique 'situation' of mortgage prisoners reveals about contemporary class inequality. *Information, Communication & Society*, 27 (10) <https://doi.org/10.1080/1369118X.2024.2334392>

Umoja Noble, S. (2018) *Algorithms of Oppression: How search engines reinforce racism*. New York, New York University Press.

Wainwright, T. (2011). Elite Knowledges: Framing Risk and the Geographies of Credit. *Environment & planning A*, 43 (3), pp.650–665.

Wallace, A. (2016) *Is there a role for US Homeownership Education in the UK?* York, Churchill Memorial Trust.

Wallace, A., Rhodes, D. and Roth, F. (2018) *Home-owners and Poverty*. York, Joseph Rowntree Foundation.

Wallace, A., Rugg, J. and Lui, J. (2022) *Do Affordable Homeownership Schemes Reduce Homeownership Risks for Lower Income Households in England?* York, Abridged Fair Finance Trust.

Webber, R. and Burrows, R. (2018) *The Predictive Postcode: The Geodemographic Classification of British Society*. Sage Publications.

Wyly, E., Moos, M., Hammel, D., and Kabahizi, E. (2009) Cartographies of race and class: Mapping the class-monopoly rents of American subprime mortgage capital. *International Journal of Urban and Regional Research* 33(2). Wiley: 332–354.

Appendix: Code Encounters project research methods

Rationale

The Code Encounters project was conceived to examine the constellation of actors that surround the digital tools used to profile tenants and mortgage borrowers in relation to their access to different housing tenures. This enabled the study to examine the full ‘regime of recognition’ (Amoore, 2020) or ‘socio-technical assemblage’ (Kitchen, 2017) connected to the production, operation and impact of the tools.

Distribution of interviews

The findings presented in the three tenure reports and associated briefings are based on the qualitative insight gleaned from 122 in-depth interviews from national stakeholders, lenders, landlords, letting agents, technology firms, social landlords, consultants, private and social housing tenants and mortgage borrowers. Table 1 shows the distribution of interviews across the different housing tenures.

Table 1 Breakdown of in-depth interviews

	Construction	Operation	Impact	Stakeholders	Total
PRS	10	13 (including 7 landlords, 5 agents and 1 insurer) (Landlord survey, n=113)	20 PRS tenants	7	50
SRS	6	15 landlords	15 SRS tenants	3	39
Lending	7 Credit tech firms	9 (including 3 brokers, 4 lenders & 2 consultants)	12 borrowers	3	31
Across Tenures				2	2
Total	23	37	47	15	122

For reference, the interview quotes in the four reports have some self explanatory labels but others are coded as followed: CR Credit risk decision software firms; SH social housing landlords; TR tenant referencing firms; ST social housing tenant; SRS firm, software firms working with social housing landlords; MB mortgage borrower; and Tenant, private rented sector tenant.

Recruitment - Technology firms constructing digital tools

The technology firms who produce the digital risk profiling tools were directly invited to participate in the research through internet searches of relevant companies or approached using snowballing techniques, where other participants recommended that we speak to firms

developing technology in this space. PRS firms were all engaged with tenant referencing and were at various points on a spectrum from hybrid analogue-digital systems to ones that were almost wholly platform-based, and reflected the whole market of providers. Social housing firms were more disparate, with some offering income maximisation and tenant onboarding tools, to some engaging with triaging and understanding the customer base, and others producing customer management systems. Credit lending firms comprised those offering digital software services to support credit risk decision-making activities including data collection and analysis or comprehensive platforms, alternative credit risk profiling for loan providers, platform mortgage broker services, and consultancy activities. There was some overlap with some lenders designing software in-house.

Recruitment - Landlords, lenders, agents and brokers using digital tools

Private sector landlords and letting agents were recruited to the study in various ways, using direct approaches after internet searches, posts inviting participation on online landlord's forums, and the online survey distributed by the National Residential Landlords Association. The online survey was hosted on the *Qualtrics* platform and obtained 113 useable responses. It asked about landlords' use of digital tools, motivations and some attitudinal questions about their sentiment towards the tools regarding accuracy, confidence, understanding etc. There were several open text boxes from which we derived qualitative data. The rest of the survey was analysed descriptively using SPSS. Landlords ranged from one large national build-to-rent operator to landlords with a single property, but were mostly those with a handful of properties, reflecting the membership base of the organisation used for recruitment.

Social landlords were recruited via direct approaches and via a research invitation circulated by the National Housing Federation, the trade body for housing associations in England. Social landlords were predominantly drawn from the north of England (n=10) and the remainder from the south (n=5), although classification is challenging as some landlords include some housing stock across multiple regions. Some were large-scale providers, others smaller community associations, but were typically medium-sized regional associations, with two local authority housing departments included.

Mortgage lenders and brokers were recruited to the study using direct approaches and snowballing techniques. The mortgage lenders included a large national lender, smaller building societies and specialist lenders, including one buy-to-let lender.

Recruitment of tenants and mortgage borrowers - the subjects of the digital tools

Private rented sector tenants were recruited with the assistance of Generation Rent, a private rented sector tenant advocacy organisation, who circulated the research invitation on our behalf. The research invitation to recruit mortgage borrowers for the study was circulated on our behalf by the Homeowners Alliance. Mortgage borrowers and private tenants were drawn from around England, although London was over-represented. Social housing tenants were identified by a market research company and drawn from London, Birmingham and Manchester in roughly equal measure. Tenants and borrowers were selected if they had moved within the last two years to aid recall of their experiences of risk-profiling and were awarded £20 Amazon vouchers

as a thank-you for their participation. Social housing tenants received £30 vouchers as they had proved harder to engage.

Analysis

Interviews were undertaken on Zoom with almost all interview audio files being professionally transcribed with the remaining three digitally transcribed, checked and corrected by the researchers. Thematic analysis was undertaken supported by Nvivo. The analysis was informed by the literature review but researchers were alive to many issues that arose directly from the data.

Ethics

Ethics approval for the research was obtained from the University of York.