


# Fairness and Discrimination in Lending Decisions: Multiple Protected Characteristics Analysis



# Financial Regulation Innovation Lab

## Who are we?

The Financial Regulation Innovation Lab (FRIL) is an industry-led collaborative research and innovation programme focused on leveraging new technologies to respond to, shape, and help evolve the future regulatory landscape in the UK and globally, helping to create new employment and business opportunities, and enabling the future talent.

FRIL provides an environment for participants to engage and collaborate on the dynamic demands of financial regulation, explore, test and experiment with new technologies, build confidence in solutions and demonstrate their ability to meet regulatory standards worldwide.

## What is Actionable Research?

FRIL will integrate academic research with an industry relevant agenda, focused on enabling knowledge on cutting-edge topics such as generative and explainable AI, advanced analytics, advanced computing, and earth-intelligent data as applied to financial regulation. The approach fosters cross sector learning to produce a series of papers, actionable recommendations and strategic plans that can be tested in the innovation environment, in collaboration across industry and regulators.

Locally-led Innovation Accelerators delivered in  
partnership with DSIT, Innovate UK and City Regions



Innovate  
UK



GLASGOW  
CITY REGION

## FRIL White Paper Series

# Fairness and Discrimination in Lending Decisions: Multiple Protected Characteristics Analysis

Kushagra Jain\*

James Bowden\*

Mark Cummins\*

\* University of Strathclyde

November 2024

**Abstract:** We build upon the comprehensive toolbox developed in Jain, Bowden and Cummins (2024), extending its applicability to multiple protected characteristics. We explore a way in which several characteristics can be simultaneously considered for multi-dimensional fairness promotion and potential mitigation of plausibly discriminatory practices. In the spirit of Jain, Bowden and Cummins (2024), once again we do this with a particular focus on US home mortgage loan applications with a granular public dataset. Finally, we address a prior deficiency, namely a worse overall model accuracy/performance as measured by Area Under the Curve (AUC). The improved AUC can be attributed to a better True Positive Rate of correctly classified loan acceptances, which is achieved with the aid of hyperparameter tuning. Specifically, we use Stratified K-Fold Cross-Validation combined with overfitting-robust hyperparameter tuning facilitated with the aid of a Grid Search. These were discussed but not explicitly implemented in the use case of Jain, Bowden and Cummins (2024). We document that even a narrow set and range of hyperparameters (mitigating the computational cost of employing the Grid Search) is sufficient to elicit these improvements. Lastly, we provide recommendations on the implications of our results including where a human-in-the-loop intervention may be merited for potentially enhancing fairness in such decision making.

# Table of Contents

1. Problem Statement.....	1
2. Use Case Demonstration.....	2
3. Conclusions, Regulatory Policy Recommendations & Future Topics.....	5
References .....	7
About the Authors .....	8

# 1. Problem Statement

The problem and motivation to study the problem we address herein is very similar to that of Jain, Bowden and Cummins (2024). However, instead of examining this from the perspective of one protected characteristic, we seek assessment of many protected characteristics at the same time. The Shapley values based explainable artificial intelligence (XAI) framework we employ is additive. Therefore, to assess the collective impact of many predictors on a decision, one could sum up their Shapley values. While a viable approach, this does not lend itself to a “human-friendly” explanation (Molnar, 2020), although it satisfies numerous other desirable XAI criteria. This is especially the case as the number of predictors increases. Specifically, “good” explanations preferred by humans, are defined as concise and single (or at most double) cause explanations, which juxtapose treatment and counterfactual groups (Molnar, 2020).

We address this issue by combining such characteristics via summation of the characteristics themselves. We subsequently create a split into two categories for several protected characteristics simultaneously based on the summed variable’s median value across all loan decisions cases. We term this dummy variable (which takes a value of 1 if the summed variable is less than or equal to the median value of the summed variable across all loan cases and 0 otherwise on this basis) a binary multiple protected characteristic. This preserves the “goodness” of the explanations elicited from our framework based on this binary multiple protected characteristic. Moreover, it enables us to apply hybrid under-over sampling in a tractable way. As noted in Jain, Bowden and Cummins (2024), applying this iteratively is problematic. In particular, it impedes the ability of the models to meaningfully predict loan outcomes, and increasingly causes imbalances for characteristics balanced in earlier iterations compared to ones balanced later, if applied this way. By reducing dimensionality to a single dummy variable accounting for multiple

characteristics, the need to iteratively balance is eliminated. Consequently, one can use the framework in Jain, Bowden and Cummins (2024) directly, once augmented in this fashion. We demonstrate this applicability with a similar use case that leverages the same data.

As our main contribution, we therefore show through our XAI analysis how potential discrimination can be identified and how potentially fairer outcomes (identified as higher misclassification of rejections by the model as acceptances) in the real-world test set can be achieved with concurrent implementation of equality of outcome and equality of opportunity scenarios in the training set. For loan applications that such a model additionally suggests should have been accepted from a fairness perspective, compared to when only equality of outcome (loan decisions) is imposed, this may not necessarily be the right outcome operationally from a credit default perspective. We recommend instead that such borderline cases should trigger a human-in-the-loop intervention at the lender and/or regulatory level and the cases re-examined.

Furthermore, another problem not fully addressed in Jain, Bowden and Cummins (2024) is that of the loss of overall model performance. Achieving potentially fairer outcomes does come at the cost of the overall accuracy (as measured by area under the curve (AUC)). Although acceptances are also better predicted by the potentially fairer model, the bias-performance trade-off worsens the overall accuracy as expected (due to increased misclassification of rejections as acceptances). To address this, we bring in an aspect discussed and built into the toolbox as a functionality, but not explicitly implemented in Jain, Bowden and Cummins (2024): hyperparameter tuning. In particular, Stratified K-Fold Cross Validation is combined with a Grid Search approach to isolate optimal hyperparameters.

Using hyperparameters is more case-specific rather than model/data/setting-agnostic as noted in Jain, Bowden and Cummins (2024).

Furthermore, it is a computationally expensive way to identify such hyperparameters. However, it has its advantages. For example, it identifies these case-specific parameters in a model/data/setting-agnostic way, finding the best possible combination of the parameters for the model within a grid of all possible combinations specified.

Coupling it with Stratified K-Fold Cross-Validation makes this approach even more comprehensive, and one might expect performance improvement as a result. Indeed, this is observed in the use case. Interestingly, even a reasonably small grid range ([0.01, 0.1, 0.3], [2, 4, 6], [0.5, 0.75, 1.0], [0.5, 0.75, 1.0] respectively) over a limited number of hyperparameters (learning\_rate, max\_depth, subsample and colsample\_bytree respectively) yields a desired accuracy improvement.

Specifically, the potentially fairer element of misclassification is retained but overall performance accuracy is improved by further enhancing the model's ability to correctly classify acceptances. This adds a different dimension to the bias-performance trade-off by tuning these hyperparameters across the range in Grid Search.<sup>1</sup> Moreover, the hyperparameters ranges used in the Grid Search are selected in a conservative way such that they are robust to overfitting concerns for the underlying models.

## 2. Use Case Demonstration

The data employed for the use case is from the publicly available US data disclosed on the Federal Financial Institutions Examination Council (FFIEC) and Consumer Financial Protection Bureau (CFPB)'s website for the Home Mortgage Disclosure Act (HMDA). As specified on their website, this is the most comprehensive source of publicly available information on the U.S. mortgage market. The

HMDA requires many financial institutions to maintain, report, and publicly disclose loan-level information about mortgages. These data help show whether lenders are serving the housing needs of their communities; they give public officials information that helps them make decisions and policies; and they shed light on lending patterns that could be discriminatory (Jain, Bowden, & Cummins, 2024).

The public data are modified to protect applicant and borrower privacy and are available for the period 2000-2023 at the time of writing. HMDA was originally enacted by Congress in 1975 and is implemented by Regulation C (Jain, Bowden, & Cummins, 2024). It captures the bulk of residential mortgage lending activity in the United States (Cortés & Strahan, 2017), and has been used in several studies and contexts for mortgages (Dlugosz, Gam, Gopalan, & Skrastins, 2023; Agarwal, Muckley, & Neelakantan, 2023; Fuster, Goldsmith-Pinkham, Ramadorai, & Walther, 2022; Bartlett, Morse, Stanton, & Wallace, 2022).

We apply the solution framework of Jain, Bowden and Cummins (2024) for a loan decision making problem (whereby the loan decision is a variable entitled *Loan Decision*, with a value of 1 if the loan is rejected and 0 if the loan is accepted). For the sake of a "human-friendly" explanation of the framework (Molnar, 2020), we focus on a single US state (Mississippi) in a single year (2018), with a single type of machine learning model (XGBoost).

Furthermore, in keeping with the qualities of a "good" explanation (Molnar, 2020), we augment our framework with a dichotomous multiple protected characteristic. We assume three plausible sources of discrimination: old age, race, and gender. Old age is captured with the variable *Old*, with a value of 1 if the applicant's age is above 62 (and 0 otherwise).

---

<sup>1</sup> Indeed, storing the relevant bias and performance metrics for each model in a Grid for comparison, or juxtaposing differences with a wider number and/or range of hyperparameters may be an interesting exercise for future studies. Moreover, comparing results from a Grid and a Random Search and assessing variability in the optimal hyperparameters across datasets would be meaningful to explore. If no significant variability exists, then the computational cost can be ameliorated. Heuristic hyperparameter values within the range of variation yielding performance improvements can be used instead with far greater computational efficiency.

Race is captured with the variable *Black*, with a value of 1 if the applicant's race is African American (and 0 otherwise). Gender is captured with the variable *Applicant\_gender\_female*, with a value of 1 if the applicant's gender is female (and 0 otherwise). To consider these together, we aggregate them via summation into a single variable. Such a variable has a maximum value of 3 and a minimum value of 0, with a higher value implying a greater degree of protected characteristics associated with a particular application. Thereafter, we create a binomial multiple protected characteristic variables based on this aggregated variable, which takes a value of 1 if the value of the aggregated variable is below the median across all loan cases, and 0 otherwise. Further, to ascertain, the impact relative to the unprotected characteristic categories in a "good/human-friendly" manner, we perform the same aggregation and dichotomization exercise for the unprotected characteristic counterparts.

The other variables used in the analysis are selected as established in the literature (Agarwal, Muckley, & Neelakantan, 2023). We apply preprocessing in the training data to balance:

1. the outcome variable (i.e. *Loan Decision*)
2. both the multinomial protected characteristic and the outcome variable

We then apply the model to the unbalanced test data (i.e. the test data subsample from the original data split into test and training subsets, without any rebalancing to correct for the class imbalance illustrated earlier) to reflect the performance and fairness of the model in a real-world pragmatic scenario.

We additionally apply hyperparameter tuning using Grid Search complemented with Stratified K-Fold Cross Validation on the training data and opt for overfitting-robust optimized parameters for better model accuracy, accounting for computation time considerations. Finally, we demonstrate two different cases and compare the Shapley values of the protected characteristic, and the performance metrics in terms of classification accuracy across the cases. As this is a single

state-year analysis, the comparison can be facilitated directly. The cases are:

- *Case i* - where preprocessing is applied only on the outcome variable, and;
- *Case ii* - where preprocessing is applied on both the outcome variable and multiple protected characteristics.

As can be seen from the Shapley value plots below in Figures 1 and 2, in both cases, the multiple protected characteristic variable *multivariate\_prot\_char\_old\_black\_female\_dichotomized* plays a role in explaining the decision making of the model. Moreover, in the second case (balancing both the protected characteristic with the outcome variable) relative to the first, its Shapley value is higher in the model's decision-making process. This evidence of the multiple protected characteristic variables being relevant is in line with Kelley, Ovchinnikov, Hardoon, & Heinrich (2022), in that feature selection that is blind to protected characteristics leads to discrimination. The second case has a better area under the curve (AUC) in terms of its performance as seen in Figure 4 compared to Figure 3: specifically, the AUC in *Case i* is 0.6717, while the AUC in *Case ii* is 0.6882. This is distinct from the use case of Jain, Bowden and Cummins (2024) where *Case i* outperformed *Case ii* (in a single protected characteristic setting) in terms of overall accuracy.

We decompose this overall model performance and assess the proportion of rejected loans misclassified by the model and the proportion of loans correctly classified by the model. For the proportion of rejected loans misclassified by the model, we observe that *Case ii's* performance (0.4587) is worse than that of *Case i's* (0.4551). If one is more interested in the proportion of loans correctly classified by the model, *Case i* (0.6896) underperforms *Case ii* (0.6986) overall. This margin is greater than in the use case of Jain, Bowden and Cummins (2024) (where it was slighter and thus the overall accuracy of *Case ii* was worse than *Case i*).

Notably, *Case ii* misclassifies more rejections as acceptances and so potentially fairer

outcomes can be said to have been achieved. This is due to the context of the rejections in this setting. Loan rejections occur if a loan application initially satisfies the approval requirements of a loan guarantor (i.e., a Government Sponsored Enterprise (GSE) – Fannie Mae or Freddie Mac – or the Federal Housing Administration (FHA)), but subsequently fails in meeting the lender’s requirements. Previous studies note that the vast majority of loans (over 90%) end up securitized by the GSEs Fannie Mae or Freddie Mac, which insure investors in the resulting mortgage-backed securities against the credit risk on the loans (Fuster, Goldsmith-Pinkham, Ramadorai, & Walther, 2022). Furthermore, these prior findings suggest that firms provide lenders with underwriting criteria that dictate whether loans are eligible for securitization and influence the pricing of loans (Fuster, Goldsmith-Pinkham, Ramadorai, & Walther, 2022). As a result, the lenders retain originated loans in portfolio (i.e. on balance sheet) and thus directly bear the risk of default for less than 10% of the loans in their sample (Fuster, Goldsmith-Pinkham, Ramadorai, & Walther, 2022).

As *Case ii* addresses the multiple protected characteristic’s class in addition to the loan outcome imbalance in *Case i*, it is therefore the correction for the former that drives the additional misclassification of rejections as acceptances. In this way, *Case ii* ensures that the feature selection is not blind to the protected characteristic and does not lead to discrimination (Kelley, Ovchinnikov, Hardoon, & Heinrich, 2022). Moreover, it is not at the cost of the correct loan acceptance classification accuracy and overall accuracy. We recommend that the additional misclassifications of rejections as acceptances in *Case ii* as opposed to *Case i* may be worth scrutinizing more closely. This is to assess if the applicants were indeed credit-worthy but were denied the mortgage due to plausible discrimination. This can be achieved with a human-in-the-loop intervention by the lender and/or regulator, reviewing these borderline instances to see if plausible discrimination drives these rejections and if they merit acceptance instead.

One needs to be cautious of approving loans merely in the name of fairness but where applications are not credit-worthy. Operational default and non-performing loan risk is exacerbated if this takes place, but an additional check in such instances may perhaps prevent unjustified loan denials. Another possibility is a sort of “micro-credit” solution (Duflo & Banerjee, 2011). Specifically, it would entail sanctioning revised loans with a lower amount for an alternate property (commensurately valued at that lower amount) if the original application was credit-worthy but not at the loan amount the applicant is seeking. From this context, coupled with less misclassification for acceptances, one may conclude *Case ii* provides overall potentially fairer and more accurate outcomes than *Case i*. Overall, this suggests that using a “*Case ii* vs *Case i*” comparative analysis, one can promote potentially fairer and more accurate outcomes with multiple protected characteristics in the data. However, what is termed potentially fairer (greater rejection misclassification as acceptance) in a purely operational sense can also be interpreted as worse performance. This is what we term a bias-performance trade-off i.e. reducing potential bias in this context is at the cost of this operational performance and is made evident from the analysis comparing the two cases.

The rationale behind the increase in multiple protected characteristic’s Shapley value ranking ties back to the intuition edified behind the rebalancing undertaken through preprocessing (i.e. hybrid over-under sampling) in Jain, Bowden and Cummins (2024). To be able to ascertain more clearly the impact of an imbalanced protected characteristic in real world decision making, this needs to be corrected for while preserving the underlying statistical properties of the data, to be able to predict potentially fairer outcomes in the unaltered test subset.



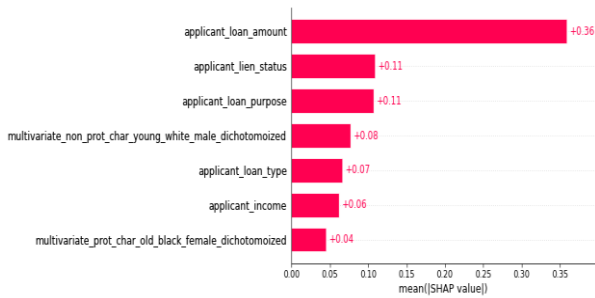


Figure 1 for *Case i*: Bar plot of mean absolute Shapley values where we rebalance only training data for outcome variable (loan decision, with a value of 1 if loan is rejected and 0 if loan is accepted)

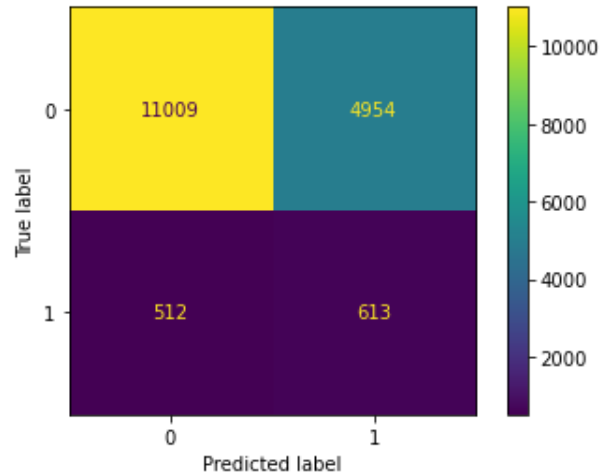


Figure 3 for *Case i*: Confusion matrix where we rebalance only training data for outcome variable (loan decision, with a value of 1 if loan is rejected and 0 if loan is accepted)

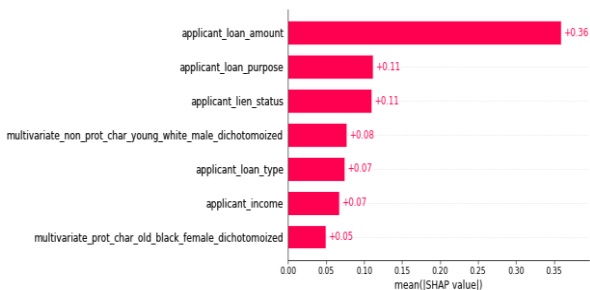


Figure 2 for *Case ii*: Bar plot of mean absolute Shapley values where we rebalance only training data for outcome variable (loan decision, with a value of 1 if loan is rejected and 0 if loan is accepted) and the multiple protected characteristic (with a value of 1 if below the median of the polynomial variable that aggregates several protected characteristics and 0 otherwise)

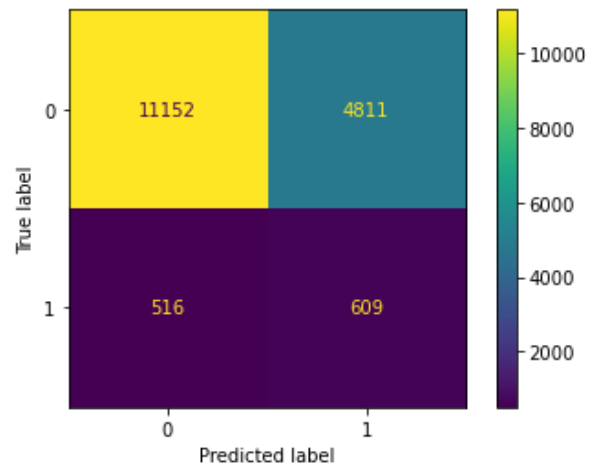


Figure 4 for *Case ii*: Confusion matrix where we rebalance only training data for outcome variable (loan decision, with a value of 1 if loan is rejected and 0 if loan is accepted) and the multiple protected characteristic (with a value of 1 if below the median of the polynomial variable that aggregates several protected characteristics and 0 otherwise)

### 3. Conclusions, Regulatory Policy Recommendations & Future Topics

We propose an extension to the framework of Jain, Bowden and Cummins (2024), examining ways in which many protected characteristics

can be simultaneously considered. We highlight an innovation to do this in two steps: first, aggregate the protected characteristics into a single compound multinomial variable, and second create a dummy multiple protected characteristic that dichotomises based on the median value of the compound variable. We further demonstrate a similar use case to that of Jain, Bowden and Cummins (2024), harnessing the same data. Specifically, we note that in order to identify plausible discrimination, it is essential to redress the class imbalance between protected characteristics in addition to the loan outcomes imbalance (Kelley, Ovchinnikov, Haroon, & Heinrich, 2022). Furthermore, by doing so and more closely examining the additional or borderline misclassifications of rejections as acceptances by redressing the class imbalance between protected characteristics, it may be possible to promote potentially fairer outcomes without harming an algorithm's overall and acceptance classification accuracy. Specifically, this may be achieved with approaches such as a human-in-the-loop intervention at a lender and/or regulator level for re-examining these cases, performing an additional credit check, or offering a "micro-credit" type of approach (Duflo & Banerjee, 2011). Finally, we mitigate the loss of overall model accuracy and performance stemming from the bias-performance trade-off. The increased misclassification of rejections as acceptances is offset by improving the correct classification of acceptances with the aid of overfitting-robust hyperparameter tuning. These hyperparameters are identified using Grid Search in conjunction with Stratified K-Fold Cross Validation. We demonstrate that this performance improvement is achievable even with a narrow set and ranges of hyperparameters, ameliorating the computational expense of finding optimal hyperparameters via Grid Search.

Possible regulatory policy implications stem from our research, which we illustrate below. In explicit terms, this may entail a closer examination of borderline cases, i.e. the cases that were misclassified by the potentially fairer model. Human intervention and reexamination

of such cases may be worth considering as a regulatory approach. This could help assess whether acceptance (rather than rejection) of such loans is indeed merited, creating a "human-in-the-loop" element to algorithmic decision making, with both algorithmic and human input considered for overall potentially fairer decisions. A further consumer duty context implication of our study is the seeking of more egalitarian outcomes for consumers within the specific setting of financial (mortgage) decisions. By identifying consumers that may be denied mortgages due to plausibly discriminatory practices, consumers and society as a whole can benefit from social upliftment and be able to afford housing of their own.

Furthermore, it may be of interest to track over time the performance of acceptances where such intervention was the cause of the acceptance, and assess performance over time with direct regulatory oversight. This could shed light on whether acceptance was desirable in the first place. Though seemingly fairer, if such acceptances subsequently result in proportionally higher non-performing assets/defaults than before, then a model that would reject such instances is a more appropriate choice. However, if such acceptances do not adversely impact the non-performing assets/defaults of lending institutions and consequently their operational risk, then employing the potentially fairer model is justifiable.

Put another way, if the potentially fairer model leads to a pareto optimal or efficient outcome in terms of operational risk and fairness, then it should be implemented. The reality may be somewhere in between: some such acceptances turn out to be defaulters while others are not. Where these actual potentially fairer loan performances lie on this operational risk spectrum may determine their feasibility; i.e. if the cost from non-performing or default instances exceeds the benefit from the profitable, non-defaulting instances, then this approach may not be operationally feasible

## References

- Agarwal, S., Muckley, C., & Neelakantan, P. (2023). Countering racial discrimination in algorithmic lending: A case for model-agnostic interpretation methods. *Economics Letters*, 111117.
- Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, 30-56.
- Buyl, M., & De Bie, T. (2024). Inherent limitations of AI fairness. *Communications of the ACM*, 67(2), 48-55.
- Cortés, K., & Strahan, P. (2017). Tracing out capital flows: How financially integrated banks respond to natural disasters. *Journal of Financial Economics*, 182–199.
- Dlugosz, J., Gam, Y. K., Gopalan, R., & Skrastins, J. (2023). Decision-Making Delegation in Banks. *Management Science*, 1-21.
- Duflo, E., & Banerjee, A. (2011). *Poor economics*. New York: PublicAffairs.
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance*, 5-47.
- Jain, K., Bowden, J., & Cummins, M. (2024, June). Promoting fairness and exploring algorithmic discrimination in financial decision making through. *FRIL White Paper Series*.
- Kelley, S., Ovchinnikov, A., Hardoon, D., & Heinrich, A. (2022). Antidiscrimination Laws, Artificial Intelligence, and Gender Bias: A Case Study in Nonmortgage Fintech Lending. *Manufacturing & Service Operations Management*, 3039–3059.
- Molnar, C. (2020). *Interpretable machine learning*. Lulu.com.

## About the Authors



**Dr James Bowden** is Lecturer in Financial Technology at the Strathclyde Business School, University of Strathclyde, where he is the programme director of the MSc Financial Technology. Prior to this, he gained experience as a Knowledge Transfer Partnership (KTP) Associate at Bangor Business School, and he has previous industry experience within the global financial index team at FTSE Russell. Dr Bowden's research focusses on different areas of financial technology (FinTech), and his published work involves the application of text analysis algorithms to financial disclosures, news reporting, and social media. More recently he has been working on projects incorporating audio analysis into existing financial text analysis models and investigating the use cases of satellite imagery for the purpose of corporate environmental monitoring. Dr Bowden has published in respected international journals, such as the European Journal of Finance, the Journal of Comparative Economics, and the Journal of International Financial Markets, Institutions and Money. He has also contributed chapters to books including "Disruptive Technology in Banking and Finance", published by Palgrave Macmillan. His commentary on financial events has previously been published in The Conversation UK, the World Economic Forum, MarketWatch and Business Insider, and he has appeared on international TV stations to discuss financial innovations such as non-fungible tokens (NFTs).

Email: [james.bowden@strath.ac.uk](mailto:james.bowden@strath.ac.uk)



**Professor Mark Cummins** is Professor of Financial Technology at the Strathclyde Business School, University of Strathclyde, where he leads the FinTech Cluster as part of the university's Technology and Innovation Zone leadership and connection into the Glasgow City Innovation District. As part of this role, he is driving collaboration between the FinTech Cluster and the other strategic clusters identified by the University of Strathclyde, in particular the Space, Quantum, and Industrial Informatics Clusters. Professor Cummins is the lead investigator at the University of Strathclyde on the newly funded (via UK Government and Glasgow City Council) Financial Regulation Innovation Lab initiative, a novel industry project under the leadership of FinTech Scotland and in collaboration with the University of Glasgow. He previously held the posts of Professor of Finance at the Dublin City University (DCU) Business School and Director of the Irish Institute of Digital Business. Professor Cummins has research interests in the following areas: financial technology (FinTech), with particular interest in Explainable AI and Generative AI; quantitative finance; energy and commodity finance; sustainable finance; model risk management. Professor Cummins has over 50 publication outputs. He has published in leading international discipline journals such as: European Journal of Operational Research; Journal of Money, Credit and Banking; Journal of Banking and Finance; Journal of Financial Markets; Journal of Empirical Finance; and International Review of Financial Analysis. Professor Cummins is co-editor of the open access Palgrave title *Disrupting Finance: Fintech and Strategy in the 21st Century*. He is also co-author of the Wiley Finance title *Handbook of Multi-Commodity Markets and Products: Structuring, Trading and Risk Management*.

Email: [mark.cummins@strath.ac.uk](mailto:mark.cummins@strath.ac.uk)



**Kushagra Jain** is a Research Associate at the Financial Regulation Innovation Lab (FRIL), University of Strathclyde. His research interests include artificial intelligence, machine learning, financial/regulatory technology, textual analysis, international finance, and risk management, among others. He is a recipient of doctoral scholarships from the Financial Mathematics and Computation Cluster (FMCC), Science Foundation Ireland (SFI), Higher Education Authority (HEA) and Michael Smurfit Graduate Business School, University College Dublin (UCD). Previously, he worked within wealth management and as a statutory auditor. He completed his doctoral studies in Finance from UCD in 2024, and obtained his MSc in Finance from UCD, his Accounting Technician accreditation from the Institute of Chartered Accountants of India and his undergraduate degree from Bangalore University. He was a FMCC Database Management Group Data Manager, Research Assistant, PhD Representative and Teaching Assistant for undergraduate, graduate and MBA programmes.

Email: [kushagra.jain@strath.ac.uk](mailto:kushagra.jain@strath.ac.uk)



Get in touch  
FRIL@FinTechScotland.com



This is subject to the terms of the  
Creative Commons license.  
A full copy of the license can be found at  
<https://creativecommons.org/licenses/by/4.0/>



University  
of Glasgow



University of  
**Strathclyde**  
Glasgow