The Gendered Algorithm:
Navigating Financial Inclusion & Equity in Access to Credit

Genevieve Smith

Kellogg College

University of Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of DPhil in the Department of International Development (ODID) at the University of Oxford

Supervisors: Professor Xiaolan Fu (ODID), Professor Ekaterina Hertog (Oxford Internet Institute (OII)), Masooda Bano (ODID)

September 2025

Word count: 93,702 (excluding references)

Table of Contents

Abstract

Artificial intelligence (AI) has the potential to help solve global problems and be employed “for good”. One area of immense recent investment and interest is the financial technology (“fintech”) sector. Boasting its ability to provide financial services for the underbanked, various startups are developing apps that collect mobile phone data and use machine learning (ML) to provide credit scores – and subsequently, opportunities to access loans – to groups often left out of traditional banking in low- and middle-income countries (LMICs).

This thesis explores whether ML-based credit assessment tools by fintech companies reinforce or mitigate gender inequitable access to finance in LMICs. I ask: (1) In what ways do the underlying logics, design choices, and management decisions of ML-based credit assessment tools by fintechs embed or challenge gender biases, and how do these choices influence gender equity in access to finance? (2) What benefits and challenges do users experience, and how do these compare between women and men? Feminist and postcolonial theory, as well as Science and Technology Studies (STS), focus on the role of power and maintain that it matters how – and by whom – technologies are designed and managed. These theories prompt my research questions, while also informing my hypotheses and parallel mixed methods approach.

My findings reveal that while fintech innovations hold promise, they fall short in addressing gender inequitable access to finance. Algorithmic lending tools are shaped by underlying logics of their developers. Developers and managers do not consider or adequately address how gender shapes access to and use of the technology, nor how gender “blind” algorithms and profit priorities can inadvertently privilege male-coded financial and digital behaviors. Perceptions of fairness by fintechs fail to challenge – and even legitimize – gender inequities in financial access. Both male and female users report positive benefits that the technology facilitates, yet gender differences persist in app access and use. In addition to the empirical contributions, my findings expand existing theories of algorithmic bias and feminist STS.

Acknowledgements

I would like to start by thanking my supervisors, Professor Xiaolan Fu, Professor Ekaterina Hertog, and Professor Masooda Bano. I was fortunate to have three supervisors, each with different and incredibly valuable perspectives and insights. I am deeply grateful to Xiaolan, whose consistent support and encouragement both pushed me and helped sustain my progress, while also enabling me to take advantage of exciting opportunities. I appreciate Masooda's perspectives and insights, which helped me to better understand and refine my arguments. I feel immense gratitude for Ekaterina who joined my supervisory team in the middle of my DPhil journey and regularly provided detailed, thoughtful comments and insights. Thank you for taking me on and believing in my potential, while being quick to share the perfect book recommendation and helping me find my footing when I felt I was stumbling. Your attention to detail and rigor has made me a better writer and scholar.

Thank you to examiners who have read my work at different DPhil milestones: Professor Laura Rival and Janaki Srinivasan. I would also like to thank Professor Janaki Srinivasan and Professor Gina Neff for reading and evaluating this work, both of whom I deeply value and respect.

I am grateful to the reading groups I've been part of at Oxford, including the Technology & Management Center for Development and the Gender and Technology group. Both provided rich intellectual spaces for me to explore, provide thoughts for others, and get feedback on my own work. Thank you to the members of these groups who read my papers and listened to my presentations, engaging with me and helping me to refine and improve my ideas.

I am also grateful to the broader intellectual communities of the Oxford Department of International Development (ODID) and the Oxford Internet Institute (OII). In particular, I would like to thank Lily Rodel and Fandi Ahmed. I was fortunate to have met Lily nearly on day one in our theory class, where it didn't take long for us to nerd out on gender and technology theory and research. Lily has provided the type of intellectual comradery and friendship I could have only hoped for. I was drawn to Fandi's smile

that lights up a room and met him in the early days of the DPhil, where we quickly bonded over a shared supervisor. I am appreciative of the extra time he spent reading my quantitative work and meeting with me to help me grow and improve.

My heart warms in thinking of the community at Kellogg College. Despite being a part-time student, I felt I belonged at Oxford because of the wonderful community at Kellogg. I will always cherish my time in Donald Michie House at the start of my doctorate, which was simply put: magical. In particular, I want to thank Adrita Haque. Fast friends right away, Adrita is quick to bring a smile to my face and made DMH feel like home. I loved exploring Oxford with you over the years.

Thank you to my research participants and to the organization that partnered with me on the survey portion of my research. I hope this thesis supports you and your work as we continue to explore the ways that AI can help us get to the world we wish to see.

Thank you to my family – Eileen, Borden, Alex, Lynda, Joe, Scottie, Karl, and Rob. Each of you supports and inspires me. Mom, thank you for the endless encouragement and enthusiasm, as well as the regular adjustments and vitamins to keep me healthy when my mind and body were tired. Dad, thank you for your love and believing in me, knowing you are proud of me means so much and pushes me to be the best version of myself. Alex, thank you for encouraging me to listen to things I don't agree with, even if I don't like it. It pushes me to better understand, reason through, and articulate my ideas and perspectives. Also, thank you for your music mixes and playlists, which have been my writing soundtracks for countless hours. Scottie and Karl, thank you for your perpetual support and for the countless meals to sustain me. Rob, thank you for the many European adventures which provided much needed breaks. I smile looking back at these memories knowing they are full of adventure, laughs, new partner skiing / snowboarding tricks, and coordinated dance moves.

Paul, where to begin… You are my number one research assistant (your title, not mine, but it is true) since the get-go. It's hard to put into words how grateful I am for all the time you spent reading sections,

brainstorming with me, and sometimes just listening. You are endlessly supportive and there for me. The depth of the care and love I feel from you and have for you has kept me going, including at times where I felt stuck or lost. Thank you for the pastries, regular chicken dinners, kiteboarding breaks, and being there, in every sense of the word.

I want to honor some special places and people. The Bodleian Library, particularly Duke Humphrey's, served as a place of constant inspiration and peace. The libraries in Oxford are a deep source of comfort and I'm grateful to have spent many hours in those walls. University parks always provided the dirt running trails, towering trees, running water, and flower patches that cleared my head and chest. I appreciate the broad AI ethics community, particularly the feminist researchers, who inspire me and shine light on important and tough questions that do not get near enough attention. The intellect and bravery I see in this community gives me energy.

Finally, I would like to thank my prior Emerging Technologies team at the US Agency for International Development (USAID). In particular, Shachee Doshi, thank you for bringing me into the team and your humble, inspiring leadership. You are not only wicked smart, but impact-driven and committed to understanding the ways that AI can and should be responsible in international development contexts. I am deeply grateful to USAID and know that work will find ways to continue, even if the organization does not. This thesis is dedicated to those who continue to envision and work towards a world that allows everyone to thrive, and the fighters holding power and technology accountable to the values of equity and justice, even when the lights feel dim.

Abbreviations

- **AI:** Artificial intelligence
- **B2B:** Business-to-business
- **B2C:** Business-to-consumer
- **CEO:** Chief Executive Officer
- **CI:** Confidence interval
- **DL:** Deep learning
- **DV:** Dependent variable
- **ECOA:** Equal Credit Opportunity Act
- **ICRW:** International Center for Research on Women
- **ICT:** Information communication technology
- **IFC:** International Finance Corporation
- **ILO:** International Labour Organization
- **IMF:** International Monetary Fund
- **IP:** Intellectual property
- **IV:** Independent variable
- **KSh:** Kenyan Shillings
- **KYC:** Know Your Customer
- **LDA:** Latent dirichlet allocation
- **LLR:** Log likelihood ratio
- **LMICs:** Low- and middle-income countries
- **LTV:** Lifetime value
- **MFI:** Microfinance institution
- **ML:** Machine learning
- **MNO:** Mobile network operator

- **MWU:** Mann-Whitney U
- **NLP:** Natural language processing
- **NGO:** Non-governmental organization
- **OECD:** The Organisation for Economic Co-operation and Development
- **OR:** Odds ratio
- **PCA:** Principal component analysis
- **SD:** Standard deviation
- **SDGs:** Sustainable Development Goals
- **SST:** Social Shaping of Technology
- **STEM:** Science, Technology, Engineering, Mathematics
- **STS:** Science and Technology Studies
- **UK:** United Kingdom
- **UN:** United Nations
- **US:** United States
- **USAID:** United States Agency for International Development
- **UX**: User experience
- **VIF**: Variance Inflation Factor

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1. Context and motivation

In recent years, artificial intelligence (AI) systems have become ubiquitous across industries and sectors, fundamentally shaping our current reality and future. AI makes it possible to automate judgements and, using machine learning (ML), learns from vast amounts of data to make decisions and predictions about people and objects related to them. While AI can reduce some aspects of human subjectivity, it can also embed and scale harmful societal biases and discrimination (Benjamin, 2019b; Z. Chen, 2023; O'Neil, 2016).

Beyond multinational companies and start-ups, AI has caught the attention of governments, multilateral agencies and non-governmental organizations (NGOs) to help solve some of the world's toughest issues and reach the Sustainable Development Goals. One area of immense recent investment and interest is the financial technology ("fintech") sector. Boasting its ability to provide financial services for the unbanked and underbanked, various startups and companies are developing apps that collect mobile phone data and use ML to provide credit scores – and subsequently, opportunities to access loans – to groups often left out of traditional banking. There is an increasing proliferation of financial technology companies ("fintechs") assessing credit and enabling access to loans through these ML-based alternative lending apps in low- and middle-income countries (LMICs) (CEGA, 2024). The prevailing rhetoric surrounding these "AI for Good" systems perpetuate a notion of "techno-solutionism" – the alluring idea that complex social problems can be solved through technology (Morozov, 2013) – while hiding under the veil of being "for good" to evade critical analysis.

Gender is a critical yet often underexplored dimension in the development and deployment of AI. Despite increasing attention to algorithmic bias, gender remains insufficiently engaged with in the research, design, and management of AI systems (D'Ignazio & Klein, 2023). This neglect could have serious implications: AI systems can both reflect and amplify existing inequalities, perpetuating or even deepening gender disparities in areas such as credit access, employment, healthcare, education, and public services. Women, particularly in LMICs, face additional challenges due to unequal access to digital tools (S. Chen et al., 2023; GSMA, 2025) and underrepresentation in tech development and management (S. M. West, 2019; E. Young et al., 2021), which can impact their use of and influence on the technologies. While ignoring gender may have unintended consequences, thoughtfully integrating and considering gender can have the potential to advance equity (D'Ignazio & Klein, 2023).

This Introduction chapter outlines the structure and background of my dissertation. First, I introduce my research questions and their significance. Subsequently, I discuss the theoretical and practical implications of my research. I then outline my dissertation in its entirety before moving to the Background chapter.

1.2. Research questions and significance

My Overarching Research Question (ORQ) is: *Do ML-based credit assessment tools by fintechs reinforce or mitigate gender inequitable access to finance in LMICs?* I have two sub-questions:

- *Research Question 1 (RQ1): In what ways do the underlying logics, design choices, and management decisions of ML-based credit assessment tools by fintechs embed or challenge gender biases, and how do these choices influence gender equity in access to finance?*
- *Research Question 2 (RQ2): What benefits and challenges do users experience in accessing and using ML-based credit assessment tools, and how do these compare between women and men?*

RQ1 examines the inputs and decision-making processes behind ML-based credit assessment tools, focusing on how organizational priorities, values, and approaches to algorithm conceptualization, design, and management shape these tools. It also explores how these decisions create affordances that can embed, deepen, or mitigate gender inequitable access to finance. RQ2 gathers gender-disaggregated insights from users to examine their experiences with ML-based credit assessment tools. To answer this second question, I present a case study of one fintech in Kenya, where these technologies are particularly popular. Grounded in my theoretical framework, this approach enabled an analysis of how individual logics, organizational ideologies, and technical design choices shape the production of ML-based credit assessment tools, while incorporating user voices to identify gender-disaggregated benefits and challenges these tools produce. Together, this provides a holistic understanding of how the technology influences gender (in)equitable access to finance. To answer my questions, I employed mixed methods utilizing qualitative data (semi-structured interviews) and quantitative data (a survey and large-scale review dataset).

While much of the overall analysis drew on qualitative methods, I had several hypotheses shaped by feminist and postcolonial theory and STS, which served as guiding expectations across my mixed methods study. They helped frame my exploration of gender dynamics in the development, management, access, and perceived impacts of ML-based credit assessment tools. First, I hypothesized that the ways these tools are conceptualized, built, and managed reinforce a gender-inequitable status quo. This expectation is informed by the theoretical understanding that technology design and management is power-laden (Irani et al., 2010), and that technology developers and managers often default to men as neutral (Wajcman, 2006), which may result in overlooking how men and women access technology differently, how historical biases in financial access are embedded in data, and how algorithmic decision-making and oversight can perpetuate inequality. This can lead to ostensibly "neutral" systems that inadvertently privilege male-coded financial and digital behaviors. Second, I hypothesized that gendered differences in access to and use of these tools contribute to persistent disparities in financial inclusion. Drawing on theory and prior research, I expected that inequalities in digital access and literacy (GSMA, 2025), as well as financial access and literacy (Hasler & Lusardi, 2017), would make these tools more accessible to men. However, I also anticipated that women who do access the tools would report greater perceived benefits, given their historically lower access to finance (World Bank, 2021) and the empowerment effects associated with access to resources (Agarwal, 1995). Taken together, these hypotheses suggested that while ML-based credit tools may offer benefits to many users, they risk reproducing gendered patterns of exclusion in access to finance. These hypotheses are not intended as fixed predictions but as theoretically grounded expectations that guided and informed both qualitative and quantitative components of my research, remaining open to empirical contestation, contextual nuance, and refinement.

This research is necessary and timely. Investors are flocking to the economic promise of massive populations beginning to do online transactions and access loans in countries where regulations are weak and many consumers remain outside formal banking systems. However, research is lacking on the social impacts of ML-based credit assessment technologies and the gender implications of delegating machine learning to assign credit scores to vulnerable populations via smartphones under for-profit fintech institutions. Investigating the gender dynamics and effects of such tools, which serve as gatekeepers to financial access, economic empowerment, and upward mobility is critical. While they may enhance financial inclusion, they can still maintain gender differences in lending and/or learn to discriminate along gendered lines. If so, the technologies can solidify gender stereotypes and norms, including expectations for women as unpaid caretakers and men as breadwinners. Over time, implications for gender (in)equality could amplify, despite providing overall reductions in the underbanked. Currently these apps are being used and deployed at scale in LMICs without strong regulations and without clear understanding of their impacts. The scenario sets the stage for an all-too-familiar "winner takes all" scenario: Investors, largely in the Global North, are enriched with little understood consequences for local communities.

1.3 Empirical, theoretical, and practical contributions

This thesis makes several empirical, theoretical, and practical contributions.

Empirically, it draws on original data to address a critical gap in the literature on ML-based credit assessment. To date, most research on ML-based credit assessment is found in computer science and economics literature, where the focus has largely been on identifying the most predictive types of machine learning models and proxy variables, and on analyzing outcomes for individuals who receive loans through these ML-based credit assessment tools. There is little research on fintechs themselves, including who is developing, deploying, and managing ML-based credit assessment technologies and how they are doing so. Addressing this gap is important: Understanding the broader impacts of these technologies requires attention to the people who design and manage them, as well as who is, or is not, getting access to loans via the tools. Relatedly, much of the previous empirical research on gender bias in AI focuses on technical aspects alone, as opposed to qualitative or mixed methods research on social causes and consequences, resulting in limited and narrow understandings (Hall & Ellis, 2023).

This thesis provides novel and critical empirical insight into the development of ML-based credit assessment tools and fintechs that develop them. Studying the actors who design and manage AI technologies is essential to uncover how and why algorithmic impacts manifest. As Barabas et al. (2020) argue, qualitative research that "studies up" – focusing on elites – is crucial for analyzing algorithmic fairness and power. Such research can reveal the root causes of algorithmic outcomes, including how data and AI models are assembled (Christin, 2020; Marda & Narayan, 2021). AI technologies are largely controlled by a relatively small group of technical elites – a "coding elite" (Burrell & Fourcade, 2021). Yet, there remains a lack of research examining this "coding elite" particularly in the financial industry. Studying elites poses well-documented challenges: elites are difficult to access, and researchers often encounter time-constraints and power asymmetries (Empson, 2018; Ma et al., 2020). These challenges are acute in studying those involved in AI system development (Collett, 2024). This difficulty may partially be linked to lack of regulation and concern from corporate stakeholders about potential regulation resulting from research exposing negative impacts of the technologies. Through interview data with fintech employees and survey data with end users obtained through a partnership with one fintech company, I am able to observe and contribute to the knowledge of how this particular technology is produced in the first place by examining organizational ideologies and design logics, while also understanding user experiences. Given the difficulty of accessing fintechs and AI developers, obtaining this data is significant and enables me to contribute new empirical knowledge about the social and institutional contexts in which algorithmic credit systems are built and deployed.

Theoretically, my findings expand existing theories of algorithmic bias and gender equity including through introducing the concept of *encoded gender norms* in Chapter 5 (my first empirical chapter), whereby data and features in algorithms are gendered, not considering the gendered nature of data and algorithms replicates and reinforces gender norms, and prioritization of inclusion can come at the cost of equity in algorithmic decision-making. Encoded gender norms offers a theoretical lens for understanding how gendered systems emerge not necessarily through explicit bias, but through the absence of critical engagement with gender, revealing how inclusion without attention to power and equity can ultimately reinforce existing gender hierarchies. My findings also expand existing theories in feminist Science and Technology Studies (STS). In Chapter 6, I introduce a new theoretical concept, *situated priorities*, which examines how priorities are constructed based on specific contexts, pressures, and epistemological standpoints, leading to varied fairness practices. This concept addresses a gap in existing theories: frameworks like Social Shaping of Technology (SST) focus on broader social and structural pressures while *situated knowledges* emphasizes epistemological positioning; yet neither fully examines how individual actors negotiate and prioritize specific outcomes within those structures that are then embedded in technology. *Situated priorities* thus offers a framing and theoretical lens for how actors' decisions are actively shaped by particular priorities (what they consider important or valuable) that are shaped by their institutional contexts and situated perspectives, leading to diverse technological outcomes.

Practically, my findings have implications for fintechs and funders. My research provides insights for fintechs and funders that seek to ensure their technologies are more gender inclusive and equitable. Recommendations for fintechs are woven into different chapters, with a comprehensive overview of recommendations captured in the Conclusion.

My research insights are also relevant for policymakers. AI is an emerging policy issue globally. Currently, economic prosperity and political power are playing leading roles in shaping laws and policies (or lack thereof) related to AI, without questioning assumptions and outcomes regarding equity and inclusion. Despite the rapid adoption and urgency, regulatory inertia persists globally in regard to AI (Zaidan & Ibrahim, 2024). My research informs regulation related to AI applications in the fintech space.

My research has two main policy audiences. The first policy audience is policymakers in LMICs. It is particularly relevant for countries with higher adoption of ML-based credit assessment, including Kenya and India, as they seek to balance innovation with fintech AI regulation. Furthermore, as the case study for my second research question is in Kenya, Kenyan policymakers may pay special attention. In May 2024, Kenya launched its *National AI Strategy 2025–2030* to position itself as a regional AI hub, emphasizing innovation, digital infrastructure, and commercialization for socioeconomic development (MICDE, 2025). In this, financial services are named as a strategic priority. The current strategy does not include legislation but points to future policy development, particularly around governance, regulation, and risk classification (Cooper et al., 2025). The strategy also lists guiding principles, which include inclusivity and non-discrimination, as well as ethical and responsible AI (MICDE, 2025). This suggests that they may be receptive to the insights and recommendations of my research. Similarly, India's *National Strategy for Artificial Intelligence #AI4All* focuses on capitalizing on and expanding AI innovation, but lacks guidance for use of AI systems and AI legislation (NITI Aayog, 2018). A 2021 report from the Indian government explains these gaps, outlines initial principles for responsible AI policy, and highlights the finance industry as a "high risk sector"; but still acknowledges policy guidance is lacking (NITI Aayog, 2021). Both of these country contexts illustrate the need for research and guidance to inform domain-specific policy, including finance and credit assessment applications.

The second policy audience is multinational organizations who are seeking to better coordinate AI ethics and international policy – including the United Nations (UN) and its Economic and Social Council, as well as the Organisation for Economic Co-operation and Development (OECD) which supports policy reforms globally for responsible AI. International organizations are recognizing the need for translational AI regulations, as aligned global regulation can reduce fragmentation and better address the global, interconnected nature of the AI industry (Zaidan & Ibrahim, 2024). Relatedly, my research informs policymakers in the United States (US) and European context as they consider how AI technologies developed and managed in the West have implications globally and impact international relations.

1.4. Definitions

There are multiple terms used throughout this thesis, which I define and justify.

I refer to the systems I'm studying as "ML-based credit assessment tools". These are tools that utilize machine learning to assess and predict one's creditworthiness to then offer a loan (or not) with certain repayment terms. In particular, I am referring to a version of these technologies that come in the form of apps that people download onto their phones. At times, I use the term "ML-based alternative lending tools". This term is often used interchangeably with "ML-based credit assessment tools," though it can also refer to the wider ecosystem of ML applications in lending beyond credit scoring itself. Machine learning is a subfield of AI. AI generally refers to "machines that respond to stimulation consistent with traditional responses from humans, given the human capacity for contemplation, judgment and intention" (D. West, 2018). In short, AI uses computers to model or replicate intelligent behavior. At the core of AI systems are algorithms (sets of mathematical instructions created by humans) that process data and inform decision-making.

Most AI applications today are powered by ML. ML involves algorithms that learn from vast amounts of data to find patterns and make predictions. ML can continually improve performance over time through the learned data (Kufel et al., 2023). Deep learning (DL) is a further subset of ML, which uses layered neural networks inspired by the human brain. DL is able to learn on its own when the model is established and improves over time (Holdsworth & Scapicchio, 2024). AI is pervasive today due to recent advances in ML and DL, as well as the explosion of available data.

Gender is a socially constructed aspect of identity that includes gender roles and norms, which can change and evolve over time and across contexts (Kaufman et al., 2023). Sex, on the other hand, refers to a person's biological characteristics. Gender is often normalized as female and male, and tends to be – though is not necessarily – aligned with sex assigned at birth. One's gender is accompanied by power dynamics that shape and maintain normalized appearances, roles, and relationships (Butler, 2011). However, neither gender or sex is binary (Kaufman et al., 2023). Gender is a spectrum, including nonbinary individuals (Kaufman et al., 2023). In this thesis, I tend to focus on implications of AI technologies particularly as it relates to women and men, including an examination of the ways that gender roles and expectations for women and men influence and are influenced by AI technology. In the survey I included an option for nonbinary, but did not receive any respondents that identify as nonbinary. My interview questions purposefully used "gender" and not "men" or "women", to allow for answers related to nonbinary people. However, all interviewees focused on gender differences related to men and women, with none discussing nonbinary people. Algorithms often encode and reinforce gender norms that are typically framed in binary terms (male, female) with significant implications for persistent and global inequalities between men and women. For this reason and given that my survey and interview data is structured around these binary categories, I focus on gendered dynamics and differences between men and women. However, this is not to suggest nonbinary people are not relevant nor affected. On the other hand, understanding how algorithmic systems interact with nonbinary identities remains a critical area for future research and one that requires special attention. In the Conclusion, I discuss the limitations of focusing only on women and men, how emerging interventions assume gender as binary (with problematic implications), and outline future research that is needed.

Gender bias can be understood as "behavior that shows favoritism toward one gender over another. Most often, gender bias is the act of favoring men and/or boys over women and/or girls" (Rothchild, 2014).

I define biased AI as AI that results in inaccurate and/or discriminatory predictions and outputs for certain subsets of the population. Gender bias in AI is therefore inaccurate and/or discriminatory predictions and/or outputs for certain gender identities (e.g., women). Building from my definition of gender bias, gender bias in AI can be seen as a form of favoritism toward one gender over another. Bias in AI encompasses two key areas. First, is performance discrepancies, including having better performance or lower accuracy for some groups (e.g., women). Second, is reinforcement or advancement of limiting stereotypes, norms, discrimination, and broader inequities. This reflects the substantial body of research studies on bias in AI, which include both these aspects (*Scientific Consensus on AI Bias*, 2025).

Gender equity refers to treating people of all genders fairly, recognizing that different genders may have different needs and face different barriers within and across various domains (Doyal, 2000). Gender equity requires paying attention to and addressing the historical and social disadvantages that prevent women and men (and other gender identities) from operating on the same level playing field. Gender equity is therefore seen as more of a *means*, with gender equality as a *result* (UNESCO, 2020). Although equality can also be referred to as treating people "equally" or the same, regardless of their identity and personal background. Equity recognizes that individuals and groups may need different resources, policies, or approaches to support pursuit of equal outcomes.

Finally, financial inclusion is typically defined as the use of formal financial services (Allen et al., 2016). Importantly, financial inclusion is not only access to financial services, but also considers the use, cost and quality of those services, including meeting consumers' needs and protecting consumers (Demirgüç-Kunt & Singer, 2017; Pesqué-Cela et al., 2021). There are different perceptions of what constitutes financial services in the context of financial inclusion. Some note that financial inclusion involves improving access to a wide variety of financial services, including insurance for example (World Bank, 2013). Others focus on reducing exclusion amongst "essential" services. Regardless, the primary aim of financial inclusion remains on enhancing the use of financial services (World Bank, 2013).

Gender equity in the context of financial inclusion requires recognizing the different ways that women, men, and other gender identities experience access, use, costs, and quality of financial services. Within this, it requires understanding and addressing the historical and systemic barriers that prevent women from equal access to and use of quality financial services. I focus on equity, rather than equality, in the context of financial services because my interest lies in how fintech interventions engage with underlying power dynamics and systemic barriers – not just in whether more people are being counted as included in financial systems or whether lending outputs are equal. Equity foregrounds the quality, context, and consequences of inclusion, emphasizing structural inequalities rather than relying on numerical access.

1.5. Chapter outlines

This thesis begins with background on financial inclusion and ML-based credit assessment tools (Chapter 2), followed by a review of the theoretical basis for my research and literature on themes and gaps in the space of AI, ML, gender, and inclusive finance (Chapter 3), and my methodology (Chapter 4). The primary chapters exploring my findings follow in Chapters 5 through 7, before concluding (Chapter 8).

Chapter 5 answers RQ1 by drawing on interview data with leaders, data scientists, and investors at fintechs developing ML-based credit assessment apps in LMICs. More specifically, it examines how the underlying logics, design choices, and management decisions of ML-based credit assessment tools by fintechs reproduce or challenge gender disparities in access to finance. Findings reveal that developers adopt "gender blind" approaches, grounded in beliefs that ML is objective and data reflects the truth. This results in a lack of grappling with the ways data, features and proxies for creditworthiness, and access to apps are gendered. While ML-based credit tools may expand financial access overall, interviewee accounts indicate that they reproduce gender inequalities: participants consistently reported that fewer women access loans and that women tend to receive smaller amounts than men. Fintechs identified demand- and supply-side reasons for gender differences, but framed them as outside their responsibility. At the same time, interviewees' observations that women are often more reliable repayers point to a potential market inefficiency and discriminatory effect – one that may be linked, alongside other factors, to profit optimization objectives embedded in system design and ongoing management. This chapter introduces the concept of *encoded gender norms*, whereby without explicit attention to the gendered nature of data and algorithmic design, AI systems reproduce existing social inequalities. In doing so, they reinforce gender norms as self-fulfilling prophecies. These findings illustrate that algorithms reflect not simply their training data, but the perspectives, priorities, and values of the people and institutions that design and deploy them.

Chapter 6 continues to answer RQ1. Using a grounded theory approach and inductive thematic analysis of the same 25 interviews, it builds on Chapter 5's findings by exploring the different ways fintechs define "fairness", including the logics therein and gender implications. More specifically, findings reveal that fintechs define fairness from both a process ("fair" algorithmic design) and outcome ("fair" credit assessment) perspective. From a process perspective, the prevalent notion of "fairness through unawareness" assumes that removing demographic data eliminates bias. However, this ignores structural inequities embedded in alternative data sources (e.g., mobile usage, digital footprints), which can act as proxies for gender and other categories such as class. From an outcome perspective, fintech actors often define fairness as benefit-maximizing, prioritizing inclusion over equity. Others equate fairness with accuracy, reducing ethical concerns to mathematical optimization. These perspectives align with logics that emphasize scalability and profit. Consequently, algorithmic lending reinforces existing gendered financial disparities while being framed as an ethical innovation. By engaging insights from industry actors, this research reveals the implicit power dynamics and cultural logics that underpin fairness narratives, including the immense influence of institutional priorities and economic logics. Rather than challenging systemic exclusion, fintechs prioritize scalability and accuracy over equity, sustaining disparities under the guise of objectivity. This is not inherently wrong, as fintechs are not acting irrationally within their market and regulatory environments. However, it highlights a need for regulation to help navigate societal questions.

Chapter 7 explores RQ2, drawing from a survey with 342 responses, coupled with a large dataset of 234,740 Google Play reviews, both of which are from users of the same fintech app in Kenya. The results illustrate gender-disaggregated insights into who is borrowing from ML-based credit assessment apps, their experiences, and resulting impacts. The findings show that while these tools expand access to credit for those excluded from formal financial systems, gender gaps persist. Women face greater barriers in navigating the app and tend to receive smaller loan amounts, while rural women are more likely to experience not receiving a loan offer compared to other women and men. Meanwhile, female entrepreneurs experience heightened impacts overall.

I conclude the dissertation with a Conclusion that includes an overview of findings, interpretations, contributions, limitations, recommendations, and a future research agenda. I revisit my overarching research question and subquestions, while summarizing my findings and providing interpretations that situate findings within broader debates about algorithmic fairness and gender equity in algorithmic credit assessment. I assess the contributions of my thesis, including by summarizing implications of the research, and examine the limitations. I provide a comprehensive overview of recommendations for fintechs, investors, and policymakers. Finally, I map out a research agenda for future work with a focus on alternative, more equitable technology futures.

CHAPTER 2: BACKGROUND

2.1. Introduction

This Background chapter begins by presenting the role of finance in poverty alleviation and women's empowerment, including a reflection on microfinance and financial inclusion in LMICs, as well as recent trends of mobile phone proliferation and fintech innovation to unlock new opportunities in access to finance. It then explores the history of credit scoring, including how credit scoring has evolved as a concept and practice globally. I present the ways that ML has entered the credit scoring space, in both developed and developing country contexts, as one of the latest fintech innovations under the "AI for good" umbrella. Finally, this Background chapter provides an overview of the different types of ML models utilized for credit assessment in LMICs.

2.2. The role of finance in poverty alleviation and women's empowerment

Development theories emphasize the key role that access to finance can play in mitigating inequality and poverty. Financial market imperfections and lack of access to finance are seen as critical reasons for persistent income inequality and reduced economic growth (Beck et al., 2009). Tackling financial market imperfections and enhancing financial inclusion can accelerate economic growth, as well as reduce income inequality and poverty; while access to finance for firms can promote entrepreneurship and innovation, resulting in firm growth and broader economic gains (Beck et al., 2009). We are far from achieving financial inclusion globally: There are 1.4 billion adults considered "unbanked" and outside of traditional financial folds (World Bank, 2021).

Despite large recent gains in financial access, persistent divides remain along axes of gender, socio-economic status, education and more. Women's account ownership is six percentage points lower than men's in LMICs (World Bank, 2021). Linked to persistent gender discrimination and limiting gender norms, women – and those with lower socio-economic status – are more likely to lack identification, lack a mobile phone, live further from formal financial services, and need more support to use financial accounts (World Bank, 2021). Women also face difficulties accessing formal finance due to lack of collateral, lack of credit history, and gender discrimination from loan officers (Demirguc-Kunt et al., 2013). Various studies illustrate gender discrimination from loan officers, including in microfinance. For example, Blanco-Oliver et al (2021) reviewed data from the World Bank's Microfinance Information Exchange Platform with records on micro-credit in 52 developing countries. The authors found persistent gender discrimination by lending officers, with less discriminatory results when borrowers and lenders are from the same sex (A. Blanco-Oliver et al., 2021). Female entrepreneurs face credit discrimination globally. A study in Spain found that female entrepreneurs are less likely to have their loan applications approved in their founding years compared to similar male peers, despite being less likely to default (de Andrés et al., 2021). This relative lack of credit access inhibits women's ability to participate and grow in self-employment opportunities, while also limiting their ability to accumulate savings, smoothen consumption in cases of income fluctuations, and deal with emergencies (Pomeranz, 2014).

Consequently, women – as well as other more marginalized groups including those from low-income backgrounds, rural residents, and the unemployed – tend to rely on more alternative lending solutions than on formal financial institutions (Kim & and Duvendack, 2024). In particular, microfinance emerged to provide access to overcome credit market failures and provide financial services to those who are largely left out of formal financial systems (J-PAL, 2023; Yunus, 2004). Microfinance grew over the years, with approximately 140 million people globally being active borrowers from a microcredit institution in 2018, an increase of 43% since 2009 (J-PAL, 2023). Reviews of randomized control trials (RCTs) on microcredit interventions have revealed mixed results, including that microcredit has not resulted in high-return investments or transformative results for borrowers (Dahal & Fiala, 2020; J-PAL, 2023; Meager, 2022). Nor has it led to broad increases in women's empowerment despite microcredit commonly targeting women (Banerjee, Duflo, et al., 2015; Pomeranz, 2014). Furthermore, some examples illustrate unintended consequences, including issues of over-indebtedness linked to providing loans to those who lacked the capacity to repay (Schicks, 2013). However, it has not been shown to cause widespread harm and may even enhance freedom in financial decision-making (Banerjee, Duflo, et al., 2015; Banerjee, Karlan, et al., 2015; Karlan & Zinman, 2011). At the same time, targeting specific groups and modifying traditional microcredit models can result in improved outcomes (Banerjee, Duflo, et al., 2015; J-PAL, 2023). As debates on microfinance continued, attention increasingly shifted toward the broader concept of financial inclusion, encompassing a range of financial services for underserved populations and highlighting mobile phones as a promising tool (Miles, 2015).

Mobile phones are increasingly key in advancing opportunities for digital payments, savings, and borrowing – a trend amplified by the COVID-19 pandemic. Worldwide financial account ownership has reached 76% of people (71% in LMICs) with mobile phones playing a critical role fueling growth in Sub-Saharan Africa and amongst women in particular (World Bank, 2021). Digital credit offers the potential for lower transactional and operational costs, along with remote accessibility, making it a promising tool for expanding financial inclusion (Bharadwaj & Suri, 2020; Björkegren & Grissen, 2018; Kim & and Duvendack, 2024).

This transformation is occurring alongside the rapid growth of the fintech sector in LMICs. Fintech firms, often startups, are leveraging digital tools to deliver services such as crowdfunding, peer-to-peer lending, and mobile-based financial products (Ha et al., 2025). The development of fintech is commonly understood in three waves: the initial spread of mobile money and digital payments; the rise of blockchain and cryptocurrencies; and, more recently, the integration of AI in lending, banking, and payment systems (Palmié et al., 2020). Together, mobile phones and fintech are reshaping the financial landscape, offering innovations for lending and access to finance, including through ML-based credit assessment. Before delving into the ML-based credit assessment space, it is important to first examine the concept of credit scoring and how it has evolved globally.

2.3. History of credit scoring and assessment

Credit scoring has roots in American society and capitalism. In his landmark book on the history of the consumer credit industry, Professor Josh Lauer (2017) traces how credit scoring emerged in the US as a tool to "democratize" access to finance and economic mobility, playing a central role in expanding commerce, enabling economic mobility, and shaping the evolution of capitalism, while being central for commerce and capitalist evolutions. It is linked to concepts and values of meritocracy, in which one can attain upward social mobility through one's own merits as opposed to one's social position and parent's socioeconomic status. Historical credit industry leaders recognized that loan repayment was more of a moral question, versus simply having the financial means. In America, a borrower's character and hard work mattered more than what one owned, which was the opposite in Europe at the time. As Josh Lauer (2017) notes, "Above all, creditworthiness is moral judgment… Credit scores are proxies for what used to be called character, reflected in the goodness or badness of one's numbers… Are we 'good' – obedient, reliable, profitable?" The formal credit industry with roots in America has long been – and continues to be – tied to morals and the notion of trustworthiness.

Early assessments were built on the "3 Cs": Character (e.g., perceived honesty, reputation), Capacity (e.g., employment, experience), and Capital (e.g., assets, liabilities). What mattered most was whether a person honored their financial obligations and made timely repayments. To predict one's character, credit professionals would explore metrics such as occupation to make inferences about individuals, or geography, including whether one was in an affluent neighborhood or not. Assessing someone's creditworthiness was seen as more of an "art", with a recognition that subjectivity and personal opinion was part of the credit assessment process. This approach did expand access to credit with millions of Americans getting access to goods and services and is linked to rising quality of life (Lauer, 2017).

A turning point came in 1958 when Bill Fair and Earl Isaac introduced computerized credit scoring, transforming credit assessment from subjective art to statistical science (Federal Reserve Board, 2007).

The credit score began as the sum of a person's numeric values on different variables related to the "3 Cs" and multiplied by their relative weights in predicting likelihood of timely repayment.

Creating numerical credit scores created a veil of objectivity to credit scoring, in which one's creditworthiness became a statistical and factual number captured in the credit score; however, the scores were far from objective. First, assigning creditworthiness and credit scores according to the "3 Cs", including one's character and morals, was not grounded in statistical theory. Psychologists contributed to the development of point-scoring systems by testing a range of variables to explore correlations between psychological traits and creditworthiness. These correlations were used to convert psychological characteristics into numerical credit scores, despite the absence of demonstrated causal relationships (Lauer, 2017). Secondly, applicants' credit histories were often ambiguous or incomplete so credit managers still would use opinion or instinct to inform how much credit one should receive. Third, information to inform credit scores contained a variety of highly personal and nonfinancial information, including intimate life details like health, legal records, and family events – all deemed fair game for assessing creditworthiness. Scoring models often used statistical correlations without theoretical grounding or causation, including seemingly arbitrary variables like car age or the first letter of a last name (Lauer, 2017). It was not clear what information would be out of bounds, with financial and personal information overlapping.

The emergence of credit bureaus further institutionalized credit scoring. Separate from conservative banks, these data-driven firms amassed vast personal datasets. Not only were nonfinancial variables incorporated, but variables that had the greatest predictive power were often nonfinancial or loosely financial (Lauer, 2017). Scoring criteria varied widely across retailers, banks, and financial companies who dealt with different populations and developed their own scorable questions and weights. "Risk" is assessed and applied unevenly in efforts to reduce losses and nonpayment (Marron, 2015). Meanwhile, the opacity of scoring systems made it difficult for individuals to understand or challenge their assessments. Computerization allowed credit markets to scale, extending credit to riskier customers by charging higher interest rates. This allowed lenders to extend loans to poorer customers in profitable ways (Lauer, 2017).

While providing credit scores and enabling access to finance allows for new access to goods and services, the history of the credit scoring industry is linked to social bias and discrimination. In the first half of the 20th century, it was standard procedure for credit managers (who were predominantly white men) to use identity markers – including race, gender, and marriage status – in assessing credit (Lauer, 2017). Risk hierarchies were attached to moral hierarchies and connected to racist and sexist assumptions about honesty and work ethic imposed by white credit managers. Even as computerization took hold in the 1960s and 1970s, women and marginalized populations continued having difficulty securing credit and loans. For women in particular, whether a woman could get credit was linked to whether they fulfilled gender norm obligations regardless of her perceived creditworthiness. More specifically, a credit score was accessible if one had a husband (Lauer, 2017).

The 1974 Equal Credit Opportunity Act (ECOA) sought to curb overt discrimination by prohibiting creditors from discriminating against credit applicants based on demographics including race, religion, nationality, sex, marital status, and age; but structural inequalities remained. Variables excluded by law were replaced with correlated proxies like occupation or homeownership, thereby perpetuating bias. Even with gender and race removed, statistical systems couldn't eliminate embedded social inequities. Predictive variables often reflected systemic disparities in employment, education, and housing. Meanwhile these decisions were embedded in pursuit of profit. As Lauer notes:

> "Consumer lenders and retailers were in business to make money… Generally speaking, they didn't reject female credit applicants because of their gender, but because women typically earned less than men and often left the workforce to have and raise children…. For individuals it was discriminatory, but not entirely arbitrary… a reflection of real structural inequalities…
>
> In theory, statistical scoring was supremely rational, consistent and transparent. In practice, discrimination was much harder to root out… Most troubling of all, variables associated with statistical credit risk were so deeply embedded in socioeconomic contexts that they were virtually impossible to disentangle. Even if gender and race were excluded from scoring systems, for example, protected classes could still experience negative bias. This occurred because gender and

> race were closely associated with 'secondary' variables that were not prohibited like occupational category, length of employment, whether one rented or owned one's home… Proxy problems immediately became apparent where legal secondary variables were intertwined with prohibited primary variables" (Lauer, 2017).

Discrimination was – and still is – not seen as inherently problematic in lending. For lenders and economists, discrimination is about rational economic choice. It is about identifying good and bad risks, which are then treated differently. From an economics perspective, discrimination is rational and variables that are statistically predictive in credit scoring models can be valid. Indeed, Fair and Isaac prioritized "information maximalism", agreeing no variable should be banned if it was statistically significant (Lauer, 2017). This idea of "information maximalism" is alive and present in data hungry machines today with nonexistent lines between credit data and noncredit data.

When ECOA passed, lenders initially worried they might be held accountable for structural inequalities due to the law's ambiguities. The law clarified that credit judgments needed to be proved "unfairly" discriminatory. Here, women denied credit at higher rates than men could be seen as discrimination; however this is not at the fault of lenders, but rather labor markets that reflected social and economic disparities in American society (Procaccini, 2015). In short, discrimination as a result of social and economic disparities was legally acceptable, as it was not the fault of lenders.

Statistical credit scoring continued to grow in America – as well as Britain – as the industry became more sophisticated. By the 1980s, credit bureau data was fully integrated into scoring models. In the 1990s, credit scores became standard for mortgage underwriting, influencing interest rates and loan terms. Credit data was increasingly monetized, sold to insurers, prospective employers, and others. Credit reporting agencies amassed massive profits conducting analysis and providing valuable market intelligence (Lauer, 2017). In recent years, machine learning emerged as a key tool in credit assessment.

Credit scoring remains controversial. At a high level, sociologists critique inherent flaws in credit scoring, including that it produces "false" needs (Marcuse, 1964) alongside expansions of personal debt (Marron, 2015). Beyond academic circles, debates around credit scoring and what kind of data should be allowed in credit scoring continue today: In June 2024, the Biden administration proposed dropping medical debt from credit reports, questioning its fairness and highlighting its disproportionate impact on marginalized groups (e.g., Black and Latino populations, those who are low income or uninsured) (Kanno-Youngs & Kliff, 2024). Credit scoring has also been criticized as risk factors don't only reflect the individual borrower, but riskiness of the environment that the borrower is in (Rice & Swesnik, 2012).

Globally, the US played a central role in exporting credit scoring systems and rules surrounding them. US-based and/or founded companies like FICO (providing scoring algorithms) and credit bureaus such as TransUnion, Experian, and Equifax (providing operating data infrastructure) played key roles in introducing and globalizing American-style credit models in emerging markets (*Credit Experts since 1899*, n.d.; Golden, 2024; Lauer, 2017). Encouraged by institutions like the World Bank and International Monetary Fund (IMF), countries adopted centralized credit bureaus to standardize and expand access to credit: By the end of 2016, approximately 134 countries of 183 countries had either a credit bureau or a credit registry, with "unprecedented growth" in emerging markets occurring since 2000 (Salamina et al., 2019). In India, for example, the first credit bureau was established in 2000 through a partnership between TransUnion and Indian banks (Reserve Bank of India, 2014). In Kenya, the first licensed bureau was launched in 2010 with TransUnion Africa playing a leading role (TransUnion Africa, 2022). In both cases, legal reforms enabled credit information sharing, and US and United Kingdom (UK)-based companies helped introduce standardized, technology-driven credit infrastructure (Mungiria & Ondabu, 2019; Reserve Bank of India, 2014). The expansions of credit bureaus and credit scoring created efficiencies, while credit bureaus also enhanced access to finance for firms (Peria et al., 2014). However, it also raised questions about the fairness and appropriateness of applying American-style models abroad (Lauer, 2017).

Following the lines of ECOA, using gender in credit scoring decision-making is prohibited by law in the majority of developed countries, as well as many LMICs (Andreeva & Matuszyk, 2019; *Gender Discrimination in Credit Access Prohibited*, 2023). While this can be important to prevent purposeful or overt discrimination, correlated proxies still encode social inequalities. Meanwhile, research on gender disparities in credit access remains limited due to lack of demographic data (Li, 2018).

Overall, the concept of creditworthiness has become firmly embedded in financial systems globally, with credit scoring today being the product of a long historical progression, in which the US has been a key driver at several milestones. This has brought various benefits, including expanding access to finance. However, tensions and challenges persist.

2.4. The integration of machine learning in financial technologies

AI has emerged as a disruptive force in the financial industry, impacting various operations in the industry (Vuković et al., 2025). This is particularly important to examine as financial technology plays a growing role in modern economies and society. The new generation of fintech (i.e. "smart fintech") increasingly leverages and integrates data science and AI to provide faster, more personalized, and scalable services across a range of financial areas such as banking, trade, lending, insurance, payments, and more (Cao et al., 2021). Indeed, the financial sector is one of the industries that has most readily adopted AI to date and continues to be a leading adopter (Chlouverakis, 2024). Financial sector spending on AI is projected to more than double to $97 billion in 2027, from 2023 (Kearns, 2023).

From credit scoring and fraud detection to customer service automation, ML is increasingly embedded in the backbone of financial platforms, evolving from rule-based automation in initial stages to increasingly sophisticated applications (Johnson et al., 2019; Vuković et al., 2025). ML models enable fintechs to extract patterns from historical and current data. They allow companies to leverage vast amounts of structured and unstructured data (i.e. data that does not follow a predefined format or structure) to identify patterns, predict outcomes, and make-real time decisions that were previously manual or rule-based. Neural networks enable systems to handle increasingly complex analytical tasks (Vuković et al., 2025). Meanwhile, advancements in generative AI technology are also opening up new use cases and opportunities, such as in customer interactions and chatbots (Chlouverakis, 2024). This integration enables fintechs to process vast amounts of data, adapt in real-time, and enhance efficiency. Taken together, ML is revolutionizing financial services globally.

Despite the rapid advancements and integration, there are immense risks alongside the performance opportunities. Research and regulation have not kept pace with the rapid adoption of AI in financial services and major research gaps remain, particularly related to ethical considerations (Vuković et al., 2025). There is a lack of research on the impacts of fintechs on financial inclusion (Ha et al., 2025). This is linked to a lack of standardized frameworks for AI implementation in financial sectors and lack of regulation (Vuković et al., 2025). There are various risks and regulatory challenges that remain including related to transparency, data privacy and security, lack of global regulatory harmonization, and algorithmic bias and fairness (Friedler et al., 2018; Vuković et al., 2025).

2.5. The rise of "AI for Good" and the promise of machine learning in financial inclusion

"AI for Good" refers to the use of AI to address social, environmental, and humanitarian challenges. This includes the use of AI to advance the UN Sustainable Development Goals (SDGs) (Tomašev et al., 2020). The field is rapidly growing as the technology advances and it is adopted by various stakeholders for social impact efforts. This includes non-profit organizations (e.g., DataKind), academic programmes (e.g., Data Science for Social good), multilateral initiatives (e.g., UN Global Pulse Labs), and corporate funders (e.g., Google AI for Good grants, Microsoft AI for Humanity). AI has gotten the attention of the UN, which is reflected in *AI for Good*, an initiative organized by the UN specialized agency for digital technologies, ITU. The initiative operates in partnership with 47 UN partners with goals of "identifying innovative AI applications, building skills and standards, and advancing partnerships to solve global challenges" (*AI for Good*, n.d.). The breadth of AI use cases across the SDGs is illustrated and captured in different repositories, including the AI for SDGs Observatory which has a catalogue of nearly 14,000 use cases across the 17 SDGs as of May 2025 (*AI for SDGs Observatory*, n.d.). In particular, use cases related to innovation and economic growth are plentiful.

AI tools to assess creditworthiness are increasingly popular in the "AI for Good" space. At a high level, the logic follows that through access to "alternative data", credit scores – and subsequently access to loans – can be provided to people left out of traditional banking folds with ensuing individual, economic, and societal benefits. The assumption is that access to finance for those deemed creditworthy by the algorithm leads to financial inclusion and related individual benefits, broader economic gains, and reduction of poverty and inequality.

These tools capitalize on the widespread use of mobile phones and internet connectivity to access alternative, non-financial data. Smartphones hold a wealth of data about an individual as they facilitate one's daily interactions and behaviors. This type of data is particularly important for people that are unbanked or underbanked, as it provides new sources of alternative, non-financial data by which to assess a person. ML models are able to parse through immense amounts of data to identify patterns predictive of credit risk, enabling the development of alternative lending tools that assign credit scores and facilitate access to loans, including for populations who were previously "credit-invisible" (OECD, 2024). These tools often operate through smartphone apps, which ask for permission to view data on a user's smartphone and collect real-time data (Björkegren et al., 2022). The application can have access to various data stored on the device (Óskarsdóttir et al., 2020). The resulting loans are typically small, carry high interest rates and have short repayment windows.

There are various benefits of ML, which make it effective and attractive for assessing creditworthiness of unbanked and underbanked people. First, ML models are able to conduct rapid, low-cost credit assessments, which is essential in dealing with smaller loan sizes that would be economically unfeasible to evaluate under more traditional, resource-intensive processes. Secondly, ML-based alternative lending tools are able to overcome infrastructural barriers that have traditionally excluded some people from formal financial systems. Unbanked people often lack collateral and credit histories resulting in a lack of information for financial institutions so lending to them is deemed riskier. ML-based credit assessment tools take advantage of alternative data that provides credit risk information and outperforms or enhances traditional credit risk assessments (Arraiz et al., 2017; Björkegren & Grissen, 2020). Furthermore, these tools help credit providers and financial institutions mitigate information asymmetry and challenges related to it, including adverse selection and moral hazard (Mhlanga, 2021). Finally, ML promises to avoid issues such as lender discrimination, which have disproportionately impacted women in LMICs.

In sum, the integration of ML and smartphone data creates opportunities for financial institutions to access and assess information, opening new pathways to financial inclusion. By generating credit scores and subsequent access to loans for unbanked and underbanked individuals, these tools promise individual, economic, and societal benefits. The core assumption is that algorithmically expanded access to credit can lead to poverty reduction and economic opportunity in ways that are fair and unbiased.

2.6. Growth of ML-based credit assessment by fintechs in LMICs and types of ML models

The growth and proliferation of ML-based credit assessment in LMICs reflects a perfect storm combination of key trends: ubiquity of smartphones and the Internet, large amounts of consumer data available, advancement in ML technology, and the adoption of credit scoring and bureaus globally. The result is that use of and investment in fintech for alternative lending is skyrocketing globally alongside excitement over its economic and social impact potential. The alternative financing market was valued at US$10.82 billion in 2022 (Grand View Research, 2023). Meanwhile the global market for AI-based credit scoring is set to experience significant growth with a projected compound annual growth rate of 25.9% from 2024 to 2031 (InsightAce Analytic, 2024). India and Kenya are two of the most common markets for these apps. In Kenya, 77% of borrowers have taken only digital loans (MicroSave, 2019) and from

2021 to 2024, people using digital credit apps and microfinance increased five-fold[1] (Central Bank of Kenya, 2024). Two of the largest fintech firms with ML-based credit assessment apps in LMICs (Branch and Tala, both headquartered in California) had over 4 million and 6 million users, respectively, with US$3.3 billion disbursed between the two firms as of 2023 (J. Robinson et al., 2023). Meanwhile, multilateral organizations, NGOs, and government agencies are eager to support and fund these types of "AI for Good" systems.

Fintech lending in LMICs took off in the mid-2010s alongside advances in AI and data availability. Early digital lenders often began with rule-based or basic statistical models, but quickly saw the potential of ML for improving credit decisions and conducting risk analysis. An early example is Safaricom's M-Shwari in Kenya, which launched in 2012 as a mobile-based microloan service and included ML in loan decisions by the mid-2010s. Mobile technology and digital payments spurred digital financial services and paved the way for ML-based lending (OECD, 2024).

Various leaders in the ML-based credit assessment space have roots in entrepreneurship and microfinance. A prominent early innovator is Branch, which is a fintech startup launched in 2015 that has since reached over 4 million customers and issued over 29 million loans (*Branch International | About*, n.d.). Branch, as well as several prominent fintechs in the ML-based credit assessment space, have leaders that spun out of Kiva (a nonprofit organization based in San Francisco, California, dedicated to expanding access to microfinance for underserved communities globally through online-facilitated loans). One interviewee compared the group of people that have left Kiva to enter the ML-based credit assessment space to the "PayPal Mafia" (a group of former PayPal employees and founders who then founded or developed other technology companies), calling them the "Kiva mafia". For example, Branch was co-founded by Matt Flannery in 2015, who was the former CEO and co-founder of Kiva, and Daniel Jung, who held senior roles in product management and business development strategy at Kiva.org. The founder of Tala (another

[1] Microfinance and digital credit apps are reported together by the Central Bank of Kenya, though future reports may differentiate the two.

prominent fintech company in the space), Shivani Siroya, previously worked in investment banking and microfinance reflecting similar paths among leaders in the growing industry.

From these beginnings, the use of ML in credit assessment grew rapidly. Fintech startups, as well as other actors such as telecom companies, realized that alternative data could be used to predict default risk when combined with ML. Today there are a variety of fintechs that operate in the ML-based credit assessment space. Many operate as B2C (business-to-consumer), while others are B2B (business-to-business), providing ML-based apps to banks, MFIs and other partners including NGOs. Across these fintechs, ML models have evolved from simple regressions to more complex algorithms that can ingest large, unstructured datasets and find patterns linked to creditworthiness.

There are different types of ML approaches deployed in fintech credit scoring, though publicly available information on specific models used by individual firms is scarce. This may be linked to the competitive nature of the market, in which algorithms are viewed as a “secret sauce” protected under Intellectual Property (IP). Nonetheless, academic literature offers insights into common methodologies and their respective strengths and limitations. Widely used models include ensemble methods (e.g., gradient boosting, random forests), logistic regression, decision trees, support vector machines, and various forms of deep learning (Dastile et al., 2020). Regardless of the different model deployed, the typical implementation process includes the following steps: (1) Data collection; (2) Data labeling; (3) Data preprocessing; (4) Model selection; (5) Model training (using the labeled data collected; unless using unsupervised learning methods); (6) Model tuning (ensuring the model works as expected and using a validation dataset to finetune the model); (7) Model testing (assess the model on new, unseen data); and (8) Model deployment (Scheffler, 2025). The selection and choice of model often depends on the desired trade-off between predictive performance, data complexity, transparency, and adaptability to different credit environments (Dastile et al., 2020).

Different models bring different benefits and limitations. Logistic regression is a widely used baseline when labeled repayment data is available and when explainability is a priority (Dastile et al., 2020). However, logistic regression methods are sensitive to data quality, are limited in handling real-world complexities, and struggle with large datasets (Zou et al., 2025). Therefore, more advanced ML techniques are increasingly utilized in credit assessment (Zou et al., 2025). This includes neural networks and support vector machines. Decision trees provide interpretability, but are prone to overfitting (Sohn & Kim, 2012). Recent credit scoring research has shifted to ensemble learning methods, with a focus on bagging (e.g., random forest) and boosting (e.g., gradient boosting). These methods can improve prediction accuracy and learn from more complex data and variables (Barboza et al., 2017; Mukhanova et al., 2024; Wang et al., 2022; Zou et al., 2025). Deep learning models have strong feature learning and non-linear mapping capabilities making them well-suited for large and unstructured datasets and attractive in credit scoring applications (Luo et al., 2017; Zou et al., 2025). However, for ensemble learning methods and, particularly, deep learning models, there are challenges in regards to interpretability and regulatory oversight. Finally, unsupervised learning methods[2], where the model is given data without labels (e.g., without default/non-default classifications), may also be used, though less frequently. For example, Principal Component Analysis (PCA) can be employed to reduce dimensionality in datasets with many variables by transforming them into a smaller set of uncorrelated components. This helps identify the most informative features while preserving as much variance as possible, making the data easier to visualize and analyze (Fan et al., 2013).

As data availability and computational power grow, research trends show a shift away from simpler models like logistic regression toward more advanced ML approaches that can extract richer patterns from large-scale, heterogeneous datasets (Dastile et al., 2020; Mukhanova et al., 2024). This is important

[2] Machine (and deep) learning can be supervised, unsupervised and reinforced. Supervised learning, which is the most common in ML-based credit assessment, includes data that is labeled to guide the pattern in identifying specific patterns. For example, a dataset might include borrower information such as income, credit score, and employment status (input data), labeled with whether each borrower ultimately defaulted on their loan (corresponding output data). Once the model learns the relationship between the input and output it can classify new data and make predictions. In unsupervised learning, the algorithm discovers patterns on its own.

as these types of models exhibit stronger performance, but are less interpretable and transparent, which means that it is more difficult to know how certain outputs are generated (Dastile et al., 2020). For this reason, ML is often referred to as “black box” technology. These issues of interpretability pose challenges in understanding and tracking bias in models. It is also important to note that ML uses pattern recognition and detects correlations, with limitations in understanding causality (Lamsaf et al., 2025).

Looking ahead, the integration of ML in fintech is likely to deepen as more data is collected from smartphones and banking apps. Fintechs are increasingly experimenting with more advanced ML approaches, such as deep learning, and integrating other forms of ML, including generative AI, in different aspects of their risk assessment processes and customer-facing applications (Chlouverakis, 2024).

CHAPTER 3: LITERATURE REVIEW AND THEORETICAL BASIS

3.1. Introduction

This chapter begins with the theoretical basis for my research questions and hypotheses. The literature review explores research on biases in ML systems, particularly gender. I then outline the thin academic research on impacts of ML-based credit assessment tools to highlight the gap my research fills. This existing research crosses a range of disciplines, particularly economics and computer science. Broadly, I demonstrate the importance of my research and outline how it aims to fill research gaps on if and how these tools reinforce or mitigate gender inequitable access to finance.

3.2. Theoretical basis

Feminist and postcolonial theory, as well as STS, focus on the role of power and maintain that it matters how – and by whom – technologies are conceptualized, designed, and managed. These theories prompt my research questions, while also informing my hypotheses and methods. Feminist and postcolonial theory, as well as STS, urge us to examine how AI tools are developed and by whom to better understand the range of ways AI technologies may reinforce bias and impact society now and over time (RQ1), and to listen to users to understand their technology experiences (RQ2). I present the overarching theoretical basis for my thesis in this section. In the chapters that follow, I further elaborate on specific components of this framework, drawing on particular theories to inform and guide the analysis as relevant to each empirical chapter's focus.

3.2.1 Feminist theory and Feminist STS

Feminist theory tells us that power, privilege, and oppression are not equally distributed (Crenshaw et al., 1996; hooks, 1984). Gender shapes the distribution of power at all levels of society, with gender subordination being a significant form of oppression across many societies globally. This subordination intersects with other forms of subordination across lines of race, class, sexual identity, nationality, ability, and more (Crenshaw et al., 1996). Systems of power reinforce and solidify privilege and oppression. Systems of power are configured and experienced in the *matrix of domination* whereby four domains of power uphold dominant groups: structural, disciplinary, hegemonic, and interpersonal (Collins, 2009). Importantly, Collins (1990) reminds us that those who have power inherently have a desire to keep that power – whether consciously or not – and can thereby reinforce inequalities that exist. The result is unjust oppression or unearned privilege related to identities, which is reflected and reinforced across the four domains of power.

Given that power is not equally distributed, we can interrogate the role of power in knowledge production and recognize that knowledge production is not objective. Forms of knowledge are *situated*, meaning they are produced by specific people in specific circumstances related to culture, history and geography (Haraway, 1988). One's *positionality* – meaning their various positions that are determined by culture and context – informs how they come to a knowledge-making process (Alcoff, 1988). Power and privilege can inform who is part of a knowledge production process and who is absent.

While knowledge production is not objective, neither is technology. Technologies are not developed passively. Rather, the Social Shaping of Technology (SST) theory posits that technology is fundamentally shaped by the social context in which it is developed, including economic, political, and social forces (MacKenzie & Wajcman, 1999). The idea that data and technologies can be objective ignores how the concept of "objectivity" has evolved over time (Daston & Galison, 2007), and that technology is inherently produced by people with perspectives and values in certain spaces and time (Kuhn, 1970; MacKenzie & Wajcman, 1999). Indeed, there is a *duality of technology* in which technology is both shaped by the people who design and build it within a specific social context, and by how people interpret, use, and assign meaning to it in everyday life (Orlikowski, 1992). In this, human agents and technology have repeated and reflective interactions that inform technology's construction and its impacts on society. Organizational and institutional arrangements also impact the ways in which technologies are developed and impact society (Fountain, 2001). Finally, it is important to note that technology as a physical object is different from technology in use, when it derives meaning and effects from contextual factors. When used, technology both influences and reflects the world around it (O'Connor & Liu, 2024). Technology is thus influenced by various factors.

Within technology development and management, gender plays an important role. Judy Wajcman (2006) introduces *technofeminsm*, in which gender is an important aspect of identity that impacts technology and is impacted by technology. She highlights "technology as a source and consequence of gender relations… [and that the] cultures and practices of technology development are gendered". Technology is essentially a cultural production tied to notions of masculinity (Wajcman, 2006). Men monopolized technology as a source of power, while women traditionally lack technological skills relative to men (Wajcman, 2002). The under-representation of women in engineering and mathematics is a legacy of historical patterns in which men dominated the profession and defined the resulting technologies (Oldenziel, 1999; Wajcman, 2002). At the same time, patriarchal structures and gender norms mutually constitute each other, being shaped by and helping to shape technology. This mutual shaping network reveals how gender and technology are co-produced (Faulkner, 2001; Wajcman, 2002).

3.1.2. Postcolonial theory and Postcolonial STS

Postcolonial theory draws attention to the historical and institutional hierarchies that continue to shape global science and technology. Technoscientific developments emerge through transnational interactions, yet dominant narratives around innovation, expertise, and intellectual property are often structured by Eurocentric assumptions that privilege Western institutions and marginalize contributions from the Global South. Indeed, patents, innovations, and science and technology narratives tend to come from the West, while entrenched hierarchies obscure the complex roles of non-Western actors in shaping technologies (Prasad, 2014). The Western/non-Western binary functions less as a geographical distinction and more as an epistemic hierarchy, which determines whose knowledge is seen as authoritative and legitimate and who controls the narratives of innovation. A postcolonial lens illustrates how science and technology are entangled with power and shaped by colonial legacies that continue to define whose knowledge is seen as legitimate.

Even the concept of numerical credit scores being applied through AI innovations has Western roots and ties to Western notions of trustworthiness, as discussed in the Background section on the history of credit scoring. While science and technology often "diffuse" to non-Western nations from the West, this diffusion is not necessarily "bad". Rather, science and technology bring important benefits including the potential for new efficiencies and opportunities, as well as greater national and commercial autonomy. Yet, this process still depends on global knowledge networks and foreign capital that reinforces positions of subordination resulting in *contradictory tendencies,* whereby those involved are both empowered and disempowered (Benjamin, 2009). Michel Foucault notes that modern Western knowledge is tied to domination, with concepts of truth, power, and knowledge operating in mutually generative ways (Foucault, 2006, 2020). Indeed, science and technology have the potential to reinforce Western perspectives, values, and morals.

Technologies – particularly when developed, managed and/or funded in the West – can reinforce colonial logics with economic and cultural benefits circulating back to the West. Postcolonial STS scholars note that sciences are best understood as *sciences of empire* (Schiebinger, 2004). Science and technology historically played a critical role in colonialism, and continue to reinforce colonial logics (Subramaniam et al, 2017). Edward Said articulates how historical "enumeration and production of statistical knowledge in colonies performed a number of functions… [including] enforcing divisions between colonial populations as a form of remote colonial rule" (Said, 1978). Enumeration was a tool of making the "Oriental" more manageable. This echoes data assemblages today, which take and use one's data to sort and supply for AI systems to understand, categorize, and manage (Adams, 2021).

In the context of financial technologies, researchers argue that fintech platforms present new forms of colonialism ("neo-colonialism") by relying on digital and data infrastructures that rely on and reproduce colonial patterns of categorization and control (Langley & Leyshon, 2022). Promises of financial inclusion can actually be forms of capitalist profiteering and exploitation, in which unstable and precarious incomes of the poor are turned into calculable risks through combining behavioral economics and predictive algorithms (Akolgo, 2023; Gabor & Brooks, 2017). Under digital financialization, poverty is not a problem to solve for, but rather a market opportunity, whereby fintechs along with development funders (and occasionally state actors) monitor and evaluate the poor as financial subjects (Jain & Gabor, 2020). The role of technology investors in neo-colonialism is key, as they exert financial power to reinforce extractive economic relations (Langley & Leyshon, 2022; Nkrumah, 1965). Meanwhile, inequalities continue to rise and data exploitation is used as a survival tool by people with precarious incomes. The focus on access to finance distracts from core economic and financial health issues, as well as solutions for more reliable jobs and stable incomes (Akolgo, 2023).

At a more granular level, it is important to consider how technologies embed and affect power hierarchies and social norms in the communities in which they operate. Indeed, science intersects with gender, race, and caste reminding us that science and society are inextricably linked (Subramaniam et al., 2017). By ignoring power hierarchies, technologies fall prey to *default discrimination* while tech products offering to fix societal issues can reproduce or deepen discriminatory processes through narrow definitions and operationalization of fairness, a *do-gooderness* that results in *technological benevolence* (Benjamin, 2019b). Ruha Benjamin (2019) captures these issues in the context of race as the "*New Jim Code"*, defined as: "the employment of new technologies that reflect and reproduce inequalities but are promoted and perceived as more objective or progressive than the discriminatory systems of a previous era". In short, by ignoring power hierarchies and social norms that reinforce them, AI tools embed those hierarchies and inequalities under veils of objectivity.

In the field of computing, postcolonial theory exposes situated knowledge and production practices, while examining uneven economic relations and cultural epistemologies. Within technology, postcoloniality recognizes that research, design, and practice are "culturally located and power laden" (Irani et al., 2010). Decolonial computing, while under-theorized, goes further by critiquing the origins of computing practices and how knowledge is produced and practiced. It emphasizes the positionality of those researching and practicing computing while focusing on decentering Euro-American centric values and universals (Hassan, 2023). Euro-American values and universals are reflected in concepts of AI ethics and accompanying normative frameworks that are almost exclusively developed in the West, but assume a set of universal concerns that can be measured and addressed objectively (Adams, 2021).

Tools could be conceptualized, built, and managed in different ways with implications towards gender equitable financial inclusion. While universal objectivity is unattainable, there remain opportunities for more inclusive knowledge production. *Strong objectivity* (Harding, 1995) can help center perspectives (*standpoints*) of those who are otherwise excluded from knowledge production processes with implications for more equitable technology development. Scholars researching decolonization of AI reiterate the importance of recentering the futures of indigenous, Black, underrepresented, and marginalized groups; whereby participation is grounded in situated knowledge and critiques of current approaches to AI innovation and ethics that can have roots in colonial histories (Hassan, 2023). Furthermore, while people in power who benefit from AI tools have the greatest incentives to reproduce the status quo, prioritizing the needs and perspectives of those at the margins may facilitate the development of systems that work for all (D'Ignazio & Klein, 2023).

3.1.3. Technology and development

Technology has played and continues to play a key role in the field of development – spanning environmental, economic, and social goals and initiatives. Amartya Sen's capability approach highlights the central importance of human capabilities in development, whereby the basic concern of human development is "our capability to lead the kind of lives we have reason to value", as opposed to economic growth (Sen, 1999). In the capability approach, the objective of development interventions should be to expand the freedom of people, with access to positive resources and the ability to make choices that matter to them (Alkire, 2005). Resources – including technologies – can be considered an input that enables capabilities, which are the basis for the achievement of "functionings" by people from which they can derive value (Haenssgen & and Ariana, 2018). Sen argues that mobile phones, a technological resource, is generally "freedom-enhancing" as the use of technology can enhance efficiency of people, while also being a tool to support liberation (Sen, 2010).

In line with the capability approach, feminist economists highlight how access to and control over resources (e.g., property, land, and credit) are tied to economic agency and women's empowerment. Naila Kabeer presents women's empowerment as the ability to make and act upon strategic life choices, with access to resources being a critical dimension in the ability to exercise choice (Kabeer, 1999). Even limited access to resources can lead to empowerment outcomes, particularly for marginalized women (Agarwal, 1995). However, in regards to financial resources, while increased formal incomes may improve women's bargaining power in the household, this does not necessarily mean women retain control of the income (Agarwal, 1995; Elson, 1999). It is critical to not only consider access to resources, including technologies, but also how women use the technologies and how their exercise of agency can be limited or policed (O'Donnell & and Sweetman, 2018). Taken together, access to resources, including technology and loans, can be impactful particularly for more marginalized women, but access to resources alone does not guarantee the enhancement of capabilities and empowerment.

The history of technology in development includes *contradictory tendencies* with benefits not equally distributed, or even harmful discriminatory tendencies exacerbated. For example, information communication technologies (ICTs) can enhance connectivity and knowledge sharing, while offering to support women's empowerment and advance economic growth (Friederici et al., 2017). However, a review of academic research on impacts of ICTs in LMICs reveal mixed and inconclusive results despite claims for self-evident positive benefits to society by tech companies and development agencies (Friederici et al., 2017). While technology can and does bring important benefits, technologies often fail to deliver on their promises due to not recognizing the role of power and hierarchies – including related to gender – that impact how benefits are distributed (D. Noble & Stalder, 1998). The use of technologies to enhance one's capabilities depends on other conversion factors (or conditions) that affect their use (Oosterlaken, 2011). Importantly, technology is not "value neutral", but depends on values that are embedded in and surround them (Oosterlaken, 2009).

There are gender gaps that exist in access and use of ICTs, as well as immense divides in the design and creation of the hardware and software underlying the ICTs. Gender and development researchers examine how gender relations are a key determinant in women's position as society, not as immutable, but as socially constructed. Women experience oppression differently, according to their ethnicity, race, class, culture, and more (Moser, 2012). This oppression intersects with technologies, which "have been designed by the few for the many, with visible and invisible power dynamics that offer context and knowledge created by elites" (O'Donnell & and Sweetman, 2018).

Development policymakers and practitioners are focused on the potential of ICTs, including AI, to deliver on economic and social development and drive the SDGs. Indeed, digital tools could help realize the SDGs, but they can also enact harm. Feminist critiques emphasize the need to move beyond considerations of the 4As (access, affordability, availability, and awareness) when it comes to technology, and incorporate questions of power and inequality in technology development and use (O'Donnell & and Sweetman, 2018; Tongia et al., 2005).

Beyond having differing access and impacts for women and men, technology impacts gender in development contexts. In a study on the impacts of ICTs on relations between sexes in urban India, Shannon Philip finds that young women and men are able to find new spaces to express themselves in new ways and form new relationships. The technologies help circumvent traditional restrictions in meeting other young people. However, Philip also finds that ICTs remain embedded with norms, prejudices, and power relations of the culture, therefore reproducing biases of patriarchal Indian society (Philip, 2018). This reminds us that ICTs and other digital technologies are not inherently good or bad, but rather impacts vary, with some impacts being positive (e.g., supporting relationship building) and others negative (e.g., reproducing patriarchal biases). They also have an influence on the practice of gender itself (O'Donnell & and Sweetman, 2018).

Technologies can be designed and used in different ways that support empowerment of marginalized communities. For example, marginalized groups can use ICTs to construct "innovative visions and practices… [and in the process of incorporating into the mainstream] also enhance impacts of these groups on constructs of power that shape our world" (Schech, 2002). It remains important not to assume positive development outcomes from proliferation of technologies, but rather recognize the role of power and examine how technologies can be conceptualized, designed, and managed differently.

3.1.4 Overall theoretical framework

Overall, my theoretical framework brings together several strands of thought. It is located in feminist STS and postcolonial STS, with foundational grounding in SST and technofeminism. These theories and perspectives shape both my research questions and methodological approach, emphasizing the co-constitutive relationship between technology and society, particularly in relation to gender. Throughout this thesis, I draw on aspects of my theoretical framework to address specific research questions.

Chapters 5 and 6 focus on RQ1: *In what ways do the underlying logics, design choices, and management decisions of ML-based credit assessment tools by fintechs embed or challenge gender biases, and how do these choices influence gender equity in access to finance?* To analyze this, I draw on Haraway's concept of situated knowledges to explore the epistemological positions and embedded logics within fintechs and ensuing design choices, including as it relates to financial inclusion and gender. I incorporate SST to examine how organizational and industry-wide structural pressures influence fintechs and inform design decisions. Additionally, I utilize technofeminism to interrogate the gendered implications of logics and decisions. This is complemented by Benjamin's concept of default discrimination, which offers a critical lens on the unintended harms and moral narratives that often accompany technological interventions.

Chapter 7 addresses RQ2: *What benefits and challenges do users experience in accessing and using ML-based credit assessment tools, and how do these compare between women and men?* Here, I employ descriptive statistics and targeted regression analysis followed by a discussion that draws on this chapter's conceptual framework. More specifically, I employ a feminist economics lens to explore how users experience impacts of the tools, with particular attention paid to how relative impacts may be greater for women (particularly from more marginalized backgrounds), while gendered tensions can persist in their access and use. I further explore how gendered experiences of the apps are linked to technological design and management choices by drawing on SST and technofeminism.

Taken together, my theoretical framework allows for a holistic assessment and informs my broader research question on whether ML-based credit assessment tools by fintechs reinforce or mitigate gender inequitable access to finance in LMICs.

3.2. Impacts of AI and ML-based credit assessment tools

I turn to existing literature on impacts of AI technologies on society, specifically focusing on how bias in AI systems disproportionately harm marginalized populations through poorer performance relative to dominant groups, replication of limiting stereotypes and norms, and/or reinforcement of discrimination and social inequities. I draw from gender, race, and STS scholars, as well as those in information studies, economics, and computer science conducting important multidisciplinary research in this realm. I then explore literature on my particular AI technology of interest: ML-based credit assessment tools. I outline the thin research on impacts of these tools, which largely sits in economics and computer science, and highlight the lack of multidisciplinary research drawing from gender and STS fields where my research fills a critical gap.

3.2.1. Bias in AI systems

Research has identified pervasive biases – which manifest in datasets, algorithms, and use of AI systems – resulting in performance discrepancies and discrimination, which illustrate that technology is not neutral (Benjamin, 2019b; Eubanks, 2018a; S. U. Noble, 2018).

Bias is often baked into datasets from which ML systems learn. Data is not objective and reflects pre-existing social and cultural stereotypes and biases (Criado-Perez, 2021; D'Ignazio & Klein, 2023). Within generative AI technologies, for example, a growing body of research examines how models reflect limiting stereotypes in training data related to gender (Girrbach et al., 2025), race (Cheong et al., 2024), and language (Fleisig et al., 2024). Further, men are seen as the "human default" resulting in ubiquitous gender data gaps (Criado-Perez, 2021). Bias can also take the form of lower accuracy or poorer performance for individuals of particular groups. These accuracy and performance discrepancies result from underrepresentation of different groups in training datasets. When a group is underrepresented, the model has fewer examples to learn from, leading to poorer performance and higher error rates for those groups compared to others that have greater representation in training data (Barocas et al., 2023). Crucially, these gaps tend to affect those already marginalized, rather than dominant or well-represented groups. A well-known study demonstrated this through an evaluation of facial recognition technologies, which performed better for white people and men, illustrating a disparity linked to the demographics represented in the training data (Buolamwini & Gebru, 2018).

Within algorithms, selection of proxies can lead to inaccurate or discriminatory predictions. For example, a widely used healthcare algorithm in the US falsely concluded that Black patients were healthier than equally sick White patients when using health costs as a proxy for health needs, ignoring that Black Americans spend less money on healthcare (Obermeyer et al., 2019). This example illustrates how inaccuracy can occur when the proxy correlates with, but does not causally determine, the outcome being predicted. In another example, a child welfare AI risk model deployed in Allegheny County, Pennsylvania, used different proxies as a measure of maltreatment, including call re-referral, which incorporates reports on abuse and neglect. However, anonymous and mandated reporters disproportionately report Black and biracial families for abuse and neglect at rates that are three and a half times higher than for white families, resulting in higher reports of maltreatment by Black and brown families (Eubanks, 2018b).

Even when algorithms do not explicitly incorporate social categories like race or gender, they can still pick up on and reproduce demographic patterns through the training data and proxy features they use, thereby masking discrimination rather than preventing it. For example, ProPublica found that an algorithm widely used by judges in the US to predict recidivism more often mislabeled white defendants as low risk, compared to Black defendants (Angwin et al., 2016). The developers used data from a questionnaire to predict one's risk, including questions such as whether you were raised by a single mother, had ever been suspended from school, or have friends and family that have been arrested. In the US, these questions reflect structural racial inequalities: only 44% of Black children grow up in two-parent households compared to 77% of white children (US Department of Justice, 2023). In practice, these questions served as proxies for race, illustrating how race is embedded in data even when the developers say their algorithms do not consider race (D'Ignazio & Klein, 2023).

Similarly, by not including social group category data, one can ignore or hide discrimination, as algorithms can pick up on statistical correlations that may be socially unacceptable or even illegal (Williams et al., 2018). A research study on admission decisions for colleges in the US found that an algorithmic prediction tool more often mis-ranked Black students when it was "race-blind" compared to when race was used (Kleinberg et al., 2018). Researchers hypothesized this difference was linked to differential access that Black students have to resources such as SAT[3] preparation, which was not accounted for in the "race-blind" tool. These cases illustrate how ignoring demographic variables, intentionally or not, can obscure the mechanisms through which bias operates, allowing discrimination to persist under the guise of neutrality.

[3] The SAT is a standardized test used for college admission in the US.

The use of ML systems can also lead to biased or discriminatory outcomes. This can occur when the system is applied in contexts, or to populations, that differ from those in which they were originally developed or trained. In these cases, the ML system may not work as well, as it does not have training data representative of the new populations and contexts (Friedman & Nissenbaum, 2017). ML tools can also be altered by their creators to purposefully result in different outcomes that can be seen as discriminatory. For example, in 2013 the New York Civil Liberties Union found that an algorithm used by the US Immigration and Customs Enforcement – which recommended whether people arrested over immigration violations should be released – had been altered. More specifically, they found that the algorithm was changed to increase detention without bond among individuals deemed "low risk". Researchers found that the tool became "increasingly impunitive" under the first Trump administration (Koulish & Evans, 2020).

These examples and the growing research field of algorithmic bias illustrates how bias in AI encompasses performance discrepancies and replication of current and historical stereotypes and inequalities. These issues are linked to technical aspects such as the underlying data, as well as choices in algorithmic design, and how it is used. Pre-existing biases, stereotypes, and inequalities are embedded – often implicitly and unintentionally – through these processes. These different forms of bias not only replicate, but can compound existing inequalities (O'Connor & Liu, 2024). A study of generative AI image models found that gender bias in model outputs was worse than gender bias that exists in society, likely due to the model learning from data reflecting existing gender bias, and then – as a pattern recognition machine – replicating this pattern at scale, thereby amplifying it (Girrbach et al, 2025). Bias in AI is not simply a technical problem. Rather, as gender bias exists in society and culture, it becomes part of the "contextual factors" influencing the development, management and use of AI.

Bias in AI can be hard to spot and diagnose. This is because, while inputs and outputs of an ML system may be known, the decision-making processes of ML algorithms (particularly those with massive amounts of data that are more complex and have higher predictive accuracy) cannot often be mapped out or understood by humans. For this reason, ML algorithms are often called “black boxes”, whereby they are incredibly difficult to decipher, even to their creators (Gryz & Rojszczak, 2021). Despite this opacity, algorithms are often seen as more fair or neutral than humans in decision-making (Gutierrez, 2021).

It matters who develops AI systems. AI systems are created in certain contexts by humans and are classification technologies (S. M. West, 2019). The perspectives, knowledge, values, and priorities of those who develop and manage AI systems will be integrated into their design and operation. People inform what AI systems are designed to optimize for and the problem AI is solving, as well as how decisions are made in dataset selection, model development, tool management, and more. The lack of diversity in STEM (Science, Technology, Engineering, Mathematics) is widely acknowledged, with those wielding power in tech being disproportionately elite, straight, white, able-bodied, cisgender men from the Global North (D’Ignazio & Klein, 2023). When it comes to AI, 22% of professionals in AI and data science fields are women – and they are more likely to occupy jobs associated with less status (E. Young et al., 2021). Within academia, 18% of authors at the leading 21 conferences are women and 80% of AI professors are men (S. M. West, 2019). There are similar disparities regarding race. A study exploring algorithmic bias and demographics of AI engineers, found that engineers from underrepresented demographic groups were more likely to recognize and work to mitigate algorithmic bias (Cowgill et al., 2020). My research is an example of how *who* is producing technology matters and impacts *how* technology is developed and ensuing impacts.

3.2.2. Existing research on impacts of ML-based credit assessment tools

Research on impacts of ML-based credit assessment in LMICs is thin and the types of research varies by discipline. The majority of existing research about the impacts of ML-based credit assessment is located in the disciplines of computer science and economics. Within computer science, research focuses on assessing the accuracy of novel models on predicting default and technical techniques to mitigate mathematical notions of bias. Economics research tends to focus on the welfare and economic impacts of those who receive loans. There is some social science research, but it tends to lack primary data and focus on broad theoretical implications and concerns.

Across existing research, several trends and research gaps emerge. Research on positive impacts of ML-based credit assessment tools for people who are underbanked primarily assesses how these tools can (a) more accurately predict loan repayment and/or (b) mitigate bias found in formal credit assessment mechanisms. Some research examines welfare impacts resulting from access to loans through ML-based alternative lending tools, including positive impacts and harms related to default and debt traps. Overall, that greater access to finance from these tools results in net positive impacts for individuals alongside broader economic growth and poverty reduction is not a given. Secondly, there is very little research on biases these tools may exhibit, including incorrectly predicting people as *not* creditworthy, as well as how these tools may reinforce or replicate patterns of economic and social inequality. In addition, there is minimal public information or research on how ML-based credit assessment tools developed by fintechs and deployed globally (rather than those created in academic or research settings) are conceptualized, designed, and managed.

3.2.2.1. Positive impacts for individuals and society

Several studies illustrate that ML-based credit assessment tools are able to accurately assess credit risk of people who are underbanked and can thereby facilitate access to finance. These studies are largely conducting research on ML tools the researchers built and assessed (as opposed to existing tools in the market). The studies show that ML tools using alternative, non-financial data can enhance credit risk assessments for underbanked people (Mhlanga, 2021). More specifically, studies examine how different types of alternative data from mobile phones increase performance of credit risk assessments for underbanked people. This includes, for example, borrowers' mobile phone history (Björkegren & Grissen, 2020); psychometric information acquired via mobile phones (Arraiz et al., 2017), and digital footprint data (i.e. accessible information related to internet usage) (Berg et al., 2020).

ML-based credit assessment tools may help control for harmful societal bias found in traditional credit scoring, and thereby enable greater access to finance for individuals and groups who experience bias and discrimination within formal financial structures. Research in the US finds that fintech tools using algorithms reduce discrimination in loans for mortgages by 40% among minoritized communities (Bartlett et al., 2019). Other research in Mexico and the Dominican Republic explores how innovations in gender-differentiated ML-based credit scoring can lead to a reduction of gender bias and discrimination found in traditional lending (CEGA, 2024; Chioda et al., 2024).

There is limited research on welfare impacts of access to loans through ML-based alternative lending tools. One randomized control study based in Nigeria examined welfare impacts of an ML-based alternative lending tool developed by a fintech company (Branch). It found that those who borrowed had a modest drop in borrowing from informal sources and an increase in financial health. The authors also found modestly positive effects of subjective well-being for borrowers (primarily through increases in mental health), but statistically insignificant on other measures of welfare (Björkegren & Grissen, 2020). Similarly, a study using data from another leading fintech providing ML-based alternative lending (Tala) in Kenya found that access to digital credit improved borrowers' financial well-being, with greater impacts among borrowers with limited access to credit and those who take loans for business purposes (A. Y. Chen et al., 2025). Research on M-Shwari in Kenya, a mobile network operator (MNO) partnering with Safaricom and the Commercial Bank of Africa to provide a credit score based on telecommunications variables, finds that their loans help households cope with unexpected income shocks (Suri et al., 2021). Research on a product from another MNO-bank partnership, Kutchova, in Malawi showed similarly modest effects (Brailovskaya et al., 2021). The small size of loans offered through ML tools can be useful for coping, but may be too small for productive investment related to human capital or business (J. Robinson et al., 2023). Only some of these studies include gender-disaggregated results and intersectionality in analyses is lacking across the studies.

3.2.2.2. Negative or neutral impacts for individuals and society

ML-based alternative lending tools offer short-term, high-interest loans that can be easy to access and result in harms related to high default rates and debt traps. One paper finds that at least 50% of Kenyan borrowers default on such loans, leading lenders to earn big profits in late fees and interest rates, while borrowers are left struggling (Qureshi, 2020). Another study illustrates that an ML-based alternative lending tool developed by a fintech company in Mexico has interest rates as high as 478% (unpublished Burlando et al (2021) from (J. Robinson et al., 2023)). Anecdotal reports note that people finance one high-interest digital loan with another, leading to scenarios of default and debt traps (J. Robinson et al., 2023). One reason is that borrowers are poorly informed about basic loan terms like late fees, while loan terms can change (Brailovskaya et al., 2021). Increasing financial literacy and transparency can be positive for both lenders and borrowers by increasing loan demand and repayment (Brailovskaya et al., 2021). While not a focus of my research, it is important to note the emergence of predatory fintech firms and apps that prey on vulnerable populations, offering loans with high interest rates and violating data privacy, resulting in immense implications including debt traps or even suicide (Singh, 2022).

Ultimately, that greater access to finance results in economic growth and poverty reduction is not a given. More borrowing is not necessarily a good thing. This is demonstrated through the anecdotal reports of debt taps related to digital loans and made clear from the US subprime mortgage crisis in 2007 to 2008 which highlighted the dangers of overborrowing and predatory lenders. Inefficient credit and loan allocation can lead to missed growth opportunities, persistent inequality and costly individual, economic, and societal disasters (Beck et al., 2009). Lack of regulation and ease of developing ML-based alternative lending apps combined with lower financial literacy in LMICs is paving the way for predatory fintech firms and accompanying dangers, with little oversight for ML-based lending overall.

This connects to a growing body of literature highlighting the colonial aspects of digital financial inclusion in LMICs. In a case study of digital financial inclusion and mobile money in Ghana, Akolgo (2023) finds that mobile money exposes users to mounting debt and high transaction costs, highlighting the extractive nature of these technologies in the Global South (Akolgo, 2023). Another case study on JUMO (a B2B fintech based in South Africa that provides ML-based credit assessment technology in various LMICs), explores how the fintech is renewing and recasting colonial relations through using tech stacks reliant on global platform ecosystems controlled by big tech (Langley & Leyshon, 2022). The authors discuss how people are reduced to a score primarily to be profitable to the developers, managers, and investors; while denying them what is actually needed for emancipation (e.g., dignified jobs and stable, secure income). Both cases illustrate how power imbalances persist under financial inclusion interventions, whereby technology is recast as a tool for categorization and extraction.

3.2.2.3. Research gaps and thin existing social science research related to bias and discrimination

Existing research largely overlooks a critical question: Who is incorrectly deemed *not* creditworthy or *less* creditworthy by the algorithm, and how does this occur? Bias in this context can manifest through performance discrepancies, where the model predicts less accurately for underrepresented or historically marginalized groups. A second concern is how these tools, regardless of predictive accuracy, may reflect and reproduce structural inequalities already embedded in society.

While some studies suggest that ML-based credit assessment tools may outperform traditional credit scoring methods and expand access to finance for previously excluded groups, research on ML models used for lending in the US also shows that these tools can still result in discriminatory outcomes. For instance, the prior discussed study by Bartlett et al. (2019) found that while algorithms reduced discrimination in mortgage lending for minoritized communities in the US, they nevertheless assigned higher interest rates to those same borrowers. This reveals how bias can persist even in supposedly improved systems. Furthermore, the researchers find that this type of harmful bias (higher interest linked to one's minoritized identity) becomes harder to spot. Similarly, Fuster et al (2021) finds that use of ML in mortgage lending in the US provides a slightly larger number of borrowers with access to credit and reduces the disparity in acceptance rates across racial and ethnic groups (i.e. there is more equal likelihood of getting approved across racial groups). However, the gap in interest rates between racial groups gets larger under ML models and there's more variation in interest rates among individuals of the same racial group, especially Black and Hispanic borrowers. Therefore, even though more people are getting access to loans, the terms of access become more unequal, particularly within and between marginalized groups (Fuster et al., 2021). The authors attribute this to ML models capturing structural differences rooted in systemic inequalities, as well as effectively inferring racial identity – even when race is not explicitly provided (Fuster et al., 2021). Relatedly, Bono et al (2021) examined ML algorithms used in credit assessments in the UK. They found that ML algorithms improve accuracy performance, but also increase gaps among groups, connecting this amplification to ML techniques being non-linear, whereby slight differences become boosted (Bono et al., 2021). These are telling studies, but there remains a lack of research on unintended consequences or discrimination in the credit domain in LMICs (Garcia et al., 2024).

To date, there is a lack of empirical research examining harmful biases in ML-based credit assessment tools in LMICs, or the socio-economic implications such biases may have for different groups and communities. While these tools may mitigate some long-standing issues in financial inclusion, such as the cognitive biases of individual loan officers that disproportionately disadvantage women, they may still produce discriminatory outcomes that result in inefficient or inequitable allocation of financial resources. Although empirical evidence remains limited, a small body of literature explores the broader concept of digital or data-driven discrimination and theorizes how it might manifest in ML-based credit tools in LMIC contexts.

Digital discrimination at a high level is defined as "discrimination in which algorithms, often based on AI techniques (such as ML), make automatic decisions that result in users treated unethically, unfairly, or just differently based on their personal data such as gender, ethnicity, religion, among others" (Criado et al., 2021). Digital discrimination can result from a combination of various data points and correlations made by the ML model, which are hard to capture and legally prove (Ferrer et al., 2021). This is further complicated as discrimination is contextually-specific, with various social and historical conditions informing discrimination and its manifestations. While digital discrimination due to ML is advancing, awareness and mechanisms to combat it are much slower. This can be linked to various reasons, such as difficulty in proving causality, opaqueness by models and developers, and the false allure that AI is unbiased (Langenbucher, 2020). Meanwhile, there remain numerous examples of discrimination in ML tools across hiring (Dastin, 2018), education (Adams et al., 2020), healthcare (Obermeyer et al., 2019), public services (Eubanks, 2018b), and more.

Social science research highlights potential biases in ML-based alternative lending tools, particularly those stemming from the data these systems use to identify patterns and make predictions about creditworthiness (see Table 1 for the main types of mobile phone data and proxies in the literature[4]). These different data can serve as proxies for common factors used in creditworthiness assessment, like capital (e.g., income) and character (e.g., one's reputation). The proxies selected for an algorithm may penalize certain identities or communities. For example, having a network of contacts deemed creditworthy can serve as a proxy that they themselves are creditworthy. Sajda Qureshi summarizes these tools as assigning "creditworthiness by association", whereby algorithms assume financially responsible people tend to be associated with other financially responsible people. Yet this can incorporate systemic biases, resulting in denying loans to people who "lack creditworthy connections" even if they are creditworthy themselves (Qureshi, 2020). In the case of tools using psychometric data, certain personality types (e.g., extroversion) serve as a proxy for a riskier borrower. Existing research does not delve into biases that can be present regarding psychological assessments for credit scoring, although there are other academic articles exploring bias in ML use for psychology assessment more broadly (Tay et al., 2022). Ultimately, primary research is lacking on how these types of data and proxies are utilized in ML-based credit assessment tools in LMICs, how they may result in different outcomes related to gender and other

[4] ML-based alternative lending tools targeting farmers augment mobile phone data with public geospatial data, satellite data, farm characteristics, and/or weather forecasts. These ML tools in developed countries may use education; job history; payday loan activities; rental records; and local public records like traffic violations and immigration status.

marginalized identities, and the socio-economic impacts for different identities based on who is deemed creditworthy (or not) by the machine. The thin existing primary research tends to be in the fields of computer science and economics, while lacking feminist STS research and perspectives.

Table 1. Mobile phone data used in ML credit assessment tools (Not exhaustive)

Data category	**Particular data tracked**	**How the data may be used**
Financial Behavior	Phone payment history; Mobile money usage; Purchase history	Proxy for willingness and ability to repay loans; may also be correlated with various features for credit assessment
Consumption behavior – Mobile phone usage	Messaging records (e.g., SMS); Call records; Email usage; Browser history	Used to infer one's behavior, character, and social networks; may also be correlated with various features for credit assessment
Consumption behavior – Habits	Online gaming activities; Online shopping activities; Social media activities; Types of apps downloaded and usage	
Mobility / Location	Geolocation	Proxy for whether one has a job (and steady income); may also be correlated with patterns of credit risk
Social networks	Phone contacts and records; Social media contacts and usage	Proxy for or correlated with riskiness and likelihood of loan repayment; ; may also be correlated with various features for credit assessment
Individual psychology and linguistic behavior	Mobile phone / online / social media behavior; Written language usage / choices / errors	Proxy for one's character (e.g., trustworthiness); linguistic data may also be correlated with default prediction
Digital footprint data	Device type and model, websites visited and whether clicked through paid ads; Time of day for purchases; Name used in email; Whether name is written in lowercase; Email errors	Proxy for income, character and reputation; may also be correlated with default prediction

While accountability remains an underexplored aspect of research on ML-based credit assessment technologies, several key concerns emerge. As ML tools tend to operate in a "black box", even developers and owners of ML tools have limited understanding of why certain predictions are made – including what variables are used and how much they are weighted in predictions (Sadok et al., 2022). Consumers may not understand why they are being denied credit or loans. Many alternative lending platforms in LMICs are legally registered as non-banking financial companies and do not legally need to explain lending decisions, leaving users with little recourse in cases of potential discrimination. While not central to my core research questions, this lack of transparency – both in the models and among the institutions that deploy them – remains a critical issue to acknowledge and I return to the implications of "black box" technology in later chapters of this thesis.

3.2.3. Gender impacts and gaps

There is little academic research on potential gender biases of ML-based alternative lending tools – whether examining performance discrepancies by gender or discrimination against particular genders. Relatedly, there is also a lack of research examining if these tools are equitably enhancing access to finance or continuing historical and current patterns of structural inequality. Indeed, prior literature reviews illustrate the lack of research related to gender bias and discrimination in AI-driven credit assessment globally (Corrales-Barquero et al., 2021). However, there are some findings to glean. First, a review of three studies on fintech firms offering digital credit finds that men are disproportionately self-selecting as potential borrowers (J. Robinson et al., 2023). The subjects included in a study on welfare impacts of those who received a loan from a fintech (Branch) were applicants who independently installed the smartphone app – 76% of whom were men (Björkegren et al., 2022), illustrating gender differences in who is accessing and using these apps. In examining welfare impacts of those who did receive a loan, the authors did not find gender differences. However, the sample size for women was smaller, and it is possible that the women who were able to access and use the app were relatively better off to begin with.

Due to gender digital divides, fewer women than men have access to smartphones and the Internet required to access apps. There are persistent global gender differences in digital inclusion and literacy (i.e. the "gender digital divide") as well as financial literacy (Hasler & Lusardi, 2017). Surveys and research by GSMA reveal that across LMICs, women are 8% less likely than men to own a mobile phone and 20% less likely to use the Internet on a mobile, with gender gaps further amplified in rural versus urban areas (Rowntree et al., 2020). In two of the most popular countries for these apps (Kenya and India) gender gaps loom: in Kenya there is a 34% gender gap in mobile Internet use, whereas in India the gender gap is 52% (Rowntree et al., 2020). Unsurprisingly, there is a fintech gender gap: in a global study of 28 countries, 21% of women use fintech products compared to 29% of men (S. Chen et al., 2023).

Gender differences in digital and financial inclusion and literacy help explain observations that women often face greater challenges using fintech apps compared to men. Linked to persistent gender discrimination and limiting gender norms, women – and those with lower socio-economic status – tend to need more support to use financial accounts (World Bank, 2021). While many women have access to the Internet and smartphones, as well as strong digital and financial literacy, the proportion of women with this access and literacy is lower as compared to men, particularly in rural areas (World Bank, 2021). Another important note is that ML tools reinforce the gender binary by not allowing for gender identities outside of male and female.

A separate thread of ongoing research explores how ML-based credit assessment tools can reduce gender financial inequality by incorporating gender into credit assessments. In most countries, incorporating demographics like gender in credit assessment is illegal given historical discrimination using such variables. Research conducted in Mexico (where it remains legal to consider gender in credit assessments) examined how incorporating gender in credit scoring algorithms can mitigate bias while not sacrificing profitability. The researchers, partnering with a large fintech lender, developed a pooled ML model (not differentiating by gender), as well as gender-segmented models, including one for women and one for men. The researchers find that 12.3% of women who would be rejected by a standard pooled ML model are approved in the gender-segmented (female) model, while only 4% of women would be approved in the pooled model and rejected in the segmented model (Chioda et al., 2024). This result implies gender bias exists in traditional ML models. It also is focused on gender as binary.

CHAPTER 4: METHODOLOGY

4.1. Introduction

This thesis presents findings from a mixed methods study exploring my Overarching Research Question (ORQ): *Do ML-based alternative lending tools by fintechs reinforce or mitigate gender inequitable access to finance in LMICs?* I have several sub-questions:

- *RQ1: In what ways do the underlying logics, design choices, and management decisions of ML-based alternative lending tools by fintechs embed or challenge gender biases, and how do these choices influence gender equity in access to finance?*
- *RQ2: What benefits and challenges do users experience in accessing and using ML-based alternative lending tools, and how do these compare between women and men?*

RQ1 explores the development of ML-based alternative lending tools, examining how organizational values, priorities, and epistemological approaches to algorithmic design and management influence their structure and operation. It also explores how these decisions create affordances that can embed, deepen, or mitigate gender inequitable access to finance. RQ2 turns to user perspectives, analyzing how men and women experience these tools similarly or differently, and how those experiences are shaped by broader structural inequalities, including related to digital tools and finance. To address these questions, I adopt a mixed-methods approach rooted in my theoretical framework. For RQ1, I conducted semi-structured interviews with individuals involved in the research, design, and management of ML-based credit assessment tools. For RQ2, I partnered with a fintech company to implement a survey of app users. I complemented this through employing computational text analysis on a large, public user review dataset of the same fintech, as well as open-ended survey responses.

My thesis uses a parallel mixed methods approach. Qualitative interviews explore the institutional and normative logics behind these systems (RQ1), while a quantitative case study investigates patterns of borrowing behavior by gender (RQ2). I follow a convergent, parallel mixed methods approach (Edmonds & Kennedy, 2017), whereby collection of qualitative and quantitative data are done simultaneously and analyzed separately. Then, in my conclusion, I compare and converge the findings to reflect on my overarching research question.

This chapter provides an overview of my analytical approach and data sources across the dissertation, while also including a reflection on positionality and ethics. Each qualitative chapter (Chapters 5 and 6) further grounds in particular aspects of my theoretical framework. Meanwhile, Chapter 7, which is the quantitative case study, provides further detail in the variable construction, conceptual framing, and quantitative methods employed. Table 2 provides an overview of the empirical chapters, summarizing their respective methods, data sources, conceptual framing, and analytical approaches.

Table 2. Overview of research approach

<table>
<tr><td colspan="6">Theoretical foundation: Feminist / postcolonial theory and Science and Technology Studies (STS)</td></tr>
<tr><td>Chapter</td><td>Focus</td><td>Method</td><td>Data</td><td>Conceptual emphasis</td><td>Analytical approach</td></tr>
<tr><td>Chapter 5: Mindsets and Management</td><td>Logics, bias, design choices</td><td rowspan="2">Qualitative (semi-structured interviews, thematic analysis)</td><td rowspan="2">Interview transcripts and recordings, memos (n=25)</td><td>Feminist STS</td><td>Abductive</td></tr>
<tr><td>Chapter 6: Perspectives and Priorities</td><td>Perceptions of fairness and linked choices</td><td>Feminist STS + algorithmic fairness</td><td>Inductive (theory-guided, open coding)</td></tr>
<tr><td>Chapter 7: New Finance, New Opportunities</td><td>Gendered borrowing patterns and impacts</td><td>Quantitative (survey and big data text analysis)</td><td>User survey responses (n=342); Google Play app store reviews (n=234,740)</td><td>Feminist STS + feminist economics</td><td>Deductive / inferential</td></tr>
</table>

4.2. Analytical approach and theoretical framework

4.2.1. Introduction to analytical approach

Given the complexity of gendered experiences with digital credit systems, and the need to understand both underlying logics of the systems (RQ1) and the lived experiences of users (RQ2), a parallel mixed methods design is well-suited for my research and to address my questions. I follow a convergent, parallel mixed methods approach (Edmonds & Kennedy, 2017), whereby collection of qualitative and quantitative data are done simultaneously and analyzed separately. In the conclusion, I compare and converge the findings to reflect on my overarching research question. This approach is both consistent with my theoretical framework, as well as my particular research questions (further discussed in the section on "Why a mixed methods approach"). I now outline the analytical approach for each question and the related chapters.

For RQ1, qualitative data are analyzed through thematic analysis to explore the underlying logics, design choices, and management decisions of ML-based alternative lending tools and how they structure the affordances of the technology. I examine the reasons fintechs offer for gender differences in lending outputs and apply socio-structural understandings to assess the role that gender norms and inequalities play in these lending differences. Coding is conducted with NVivo software using inductive and deductive approaches.

Chapter 5 delves into the results of this analysis, outlining how the underlying logics, design choices, and management decisions of these tools embed or challenge gender biases. Through this analysis, fairness emerged as a central theme, with perceptions of fairness often reinforcing gender "blind" approaches and legitimizing gender differences in algorithmic-facilitated lending. Chapter 6 therefore continues from Chapter 5, building on this theme and engaging more deeply with the concept of algorithmic fairness. This reflects a flexible, iterative approach consistent with interpretative qualitative research (Charmaz, 2024; Denzin & Lincoln, 2017).

For RQ2, I conducted a case study of one fintech company in Kenya, drawing on two data sources: original survey data and a large public dataset of user reviews that I scraped from Google Play. My primary analysis focuses on the survey data, using a multi-method, staged approach integrating descriptive statistics, regression modeling, and mediation analysis to investigate gendered experiences of ML-based lending in Kenya. The analysis was grounded in the theoretical expectation and my hypotheses that women face more pronounced challenges in accessing and using ML-based credit assessment apps, while also potentially experiencing greater benefits. I also employed computational social science methods, specifically text analysis, on the large user review dataset, which offered unstructured data for complementary insights and enabled me to compare patterns observed in the survey findings. While not intended as formal triangulation, the review data provided an additional lens through which to assess whether themes identified in the survey data are echoed in users' unsolicited accounts via reviews on a public platform. I additionally conducted text analysis on open-ended survey responses to contextualize and deepen interpretation of the structured survey results. Finally, I include a set of exploratory findings from the survey data that provide context on app users and their loan behaviors.

This approach – complementing survey data with a large, unstructured dataset of publicly available user reviews – aligns with the emerging field of computational social science, which uses digital trace data to study social patterns, behaviors, and discourse at scale (Freelon, 2014; Golder & Macy, 2014). Computational text analysis methods including sentiment analysis and topic modeling are well-established tools for examining large, unstructured corpora and offer a valuable complement for theory-driven, mixed methods research by enabling systematic identification of latent themes (Chakrabarti & Frye, 2017). Text analysis has also emerged as an area for algorithmic auditing and assessing algorithmic bias (Schwemmer et al., 2020). While computational approaches come with concerns, including underlying bias in the datasets and models employed (De et al., 2023), I mitigate this through a theory-driven, hypothesis-guided design where text analysis is used as a complement and triangulation strategy. Finally, in line with the emerging area of AI "incident databases" (which are databases in which users can share incidents and experiences with AI technologies) (McGregor et al., 2022), user reviews can offer a valuable lens into user-reported harms and experiences, allowing me to surface both individual and structural patterns.

4.2.2. Theoretical framework and chapter-specific lenses

Feminist and postcolonial theory, as well as STS, focus on the role of power and maintain that it matters how – and by whom – technologies are conceptualized, designed, and managed. These theories prompt my research questions, while also informing my hypotheses and methods, including the orientation "upwards" to study developers and managers of AI systems. Feminist and postcolonial theory, as well as STS, urge us to examine how AI tools are developed and by whom to better understand the range of ways AI technologies may reinforce power dynamics and impact society now and over time (RQ1), while also encouraging me to listen to voices of users themselves (RQ2) to understand the benefits and challenges users experience in using the technology.

My theoretical basis, which is laid out in Chapter 3, particularly draws on *situated knowledges* (Haraway, 1988) and technofeminism (Wajcman, 2006), to recognize the role of power in knowledge and technology production. In this, forms of knowledge are *situated*, meaning they are produced by specific people in specific circumstances related to culture, history, and geography (Haraway, 1988). One's *positionality* – meaning their various positions that are determined by culture and context – informs how they come to a knowledge-making process (Alcoff, 1988). Power and privilege can inform who is part of a knowledge production process and who is absent. While knowledge production is not objective, neither is technology. Technologies are not developed passively (MacKenzie & Wajcman, 1999). Technofeminsm highlights that gender in particular is an important aspect of identity that impacts technology production, while being impacted by implementation and adoption of technology (Wajcman, 2006). Meanwhile, postcolonial theory highlights how science and technology patents and narratives tend to come from the West (Prasad, 2014). Even the concept of numerical credit scores being applied through AI innovations has Western roots and ties to Western notions of trustworthiness (Lauer, 2017).

I also draw from Ruha Benjamin's concept of *default discrimination*, which illustrates how technologies perpetuate discrimination by replicating inequities that exist in society. Meanwhile, tech products offering to fix societal issues can reproduce or deepen discriminatory processes through narrow definitions and operationalization of fairness resulting in *technological benevolence* (Benjamin, 2019b). This concept is particularly important in the international development field as "AI for Good" is increasingly deployed across a range of sectors and contexts in LMICs. While technology can and does bring important benefits, technologies often fail to deliver on their promises due to not recognizing the role of power and hierarchies – including related to gender – that impact how benefits are distributed (D. Noble & Stalder, 1998). Taken together, my theoretical basis illustrates that positive outcomes cannot be assumed from the proliferation of AI; rather, there are ways that power manifests in technology development and management, with implications for marginalized communities. It guides my research questions and mixed methods approach.

While the dissertation is broadly grounded by feminist and postcolonial STS, each empirical chapter draws on a specific conceptual framework informed by its analytical approach and data that is elaborated in each chapter. The qualitative chapters (Chapters 5 and 6) draw more directly on technofeminism and STS to examine the underlying logics and choices of developers and managers. Meanwhile, the quantitative chapter (Chapter 7) employs a conceptual framework that integrates feminist economics, gender and development research, and technofeminism to inform my hypotheses and guide analysis of gendered differences in the benefits and challenges users experience in accessing and using ML-based credit assessment tools. This particular conceptual framework is further detailed in Chapter 7.

4.2.3. Why do I employ a mixed methods approach?

The mixed methods design is consistent with my theoretical foundations in feminist theory and STS. At a high level, these theoretical and disciplinary foundations call for attention to both underlying systems that STS emphasizes, and the voices of people themselves that feminist research urges. They also hold that different methods produce different kinds of knowledge (Haraway, 1988), and together offer a more holistic understanding of gendered realities and implications of technologies. Indeed, feminist researchers embrace multiplicity in research methods (Reinharz, 1992), employing mixed methods to address complex research questions and develop more nuanced understandings of lived experiences (Hesse-Biber et al., 2007; Hesse-Biber & Griffin, 2015).

More specifically, qualitative research is appropriate for my first question, and quantitative for the second, with both then feeding into my high-level question. For RQ1, qualitative data allows me to examine why and how impacts of the ML tools can manifest, including as they relate to underlying logics, as well as internal organizational structures, priorities, and approaches to help illuminate the "ground truth" around root causes of how data and AI models are assembled (Marda & Narayan, 2021). Relatedly, as supported in interpretive research theory, properly identifying actions – or design choices in the context of algorithms – requires grasping relevant beliefs and meanings of a person (Bevir & Blakely, 2018). Ultimately, qualitative research allows for better understanding of cultural and structural factors shaping algorithms, while quantitative methods are limited in asking questions of power hierarchies and asymmetries (Christin, 2020). Meanwhile, companies can be resistant to providing data on their ML models due to business risk and confidentiality. I therefore employ elite interviewing and "studying up" which allow me to dissect the logics and decisions of fintechs (Nader, 1972). Decisions made by small groups of people at largely startup companies impact access to finance and economic opportunities for millions of people globally. In line with my theoretical basis, the work of data science and algorithms is a sociotechnical process. Too often the assumptions and premises of how technology is conceptualized, designed, and managed remained unexamined, with research on fairness and bias having a narrow focus on quantitative analysis that misses key understandings of the ways power manifests in and through technology. Studying up is a theoretical orientation that allows for a reorientation “upward” that is increasingly recognized as critical in studies of algorithmic fairness and power (Barabas et al., 2020). This approach is not without challenges, as tech companies do not need to share information about their models and ways of thinking, while there are also considerations around intellectual property in highly competitive industries such as algorithmic-facilitated lending.

For RQ2, quantitative data enables a deep dive in gender-differentiated user experiences, focusing on Kenya as a single-country case study. Combining survey research with computational analysis of a large user review dataset allows me to engage a large and diverse sample and extract real-world insights into how users interact with ML-based credit apps. My analysis captures both user behaviors and perceptions, while also illuminating how the design and functionality of the app may shape patterns of access and use. Analyzing structured survey data, and then complementing this through analysis of unstructured user-generated reviews enables me to cross-validate key themes and conduct a more comprehensive analysis than what could be provided through a single data source.

4.3. Methods and data sources

In this section I delve into the specific methods and data sources of my two research designs. I begin with RQ1, which employs qualitative data and specifically, semi-structured interviews with fintechs. I outline my sampling frame and approach, as well as details of my analysis. I then explore RQ2, which draws on quantitative data, including survey data and a large user review dataset. For the quantitative data, I provide an overview here, while further detail on the data, my variables, and the empirical and analytical approach is provided in Chapter 7.

4.3.1. RQ1: Semi-structured interviews

My first research question asks: *In what ways do the underlying logics, design choices, and management decisions of ML-based alternative lending tools by fintechs embed or challenge gender biases, and how do these choices influence gender equity in access to finance?* To answer this question, I conducted 25 semi-structured interviews with representatives or partners of 18 companies developing and managing ML-based alternative lending apps. Inclusion criteria for the interviewees is detailed in Table 3. In particular, I sought to interview corporate leaders and data scientists of fintechs, as well as investors and research partners. In my sample I include corporate leaders and investors as they are best placed to comment on organizational priorities, perspectives, and values, and how those inform algorithmic decisions and management. I include data scientists as they are best placed to throw light on the particular ways ML-based credit assessment tools were developed and choices made in technology design. Research partners provide richer understanding of the models and how impacts are understood and assessed to date. All types of interviewees are well placed to discuss actual and perceived impacts of the technologies, including as it relates to gender, and support a rich analysis.

One of my interview participant criteria was their ability to speak English. This decision was both practical and reflective of my own linguistic capacities and positionality. In my landscape review, all fintech companies identified had English-language websites, and those I contacted communicated in English. While this may have limited access to companies operating exclusively in other languages (e.g., Mandarin), I did not encounter cases where a company of interest was excluded due to language. All individuals who agreed to participate were able to conduct the interview in English (even though it was not the first language for several participants), suggesting that this criterion did not significantly constrain the interviewee pool.

As an important caveat, within the ML-based credit assessment industry there are various types of companies. I focus on fintechs that leverage apps and use alternative data to inform credit assessment in

LMICs. I do not include predatory lending companies, as they are not a focus of my research. Nor do I include telecom companies that use data from smartphones, largely financial-related, to provide micro-loans. I focus on this segment of ML-based credit assessment lenders given their roots with Silicon Valley, their mission-focused approaches on financial inclusion, their increasing prominence globally (including in "AI for Good" spaces), and the lack of research on them including implications related to financial inclusion and gender.

Table 3. Interview participants criteria

	Criteria
Inclusion criteria	● Employee (current or previous, either management leader and/or data scientist), investor / funder or partner researcher of a fintech organization that develops and/or manages an ML-based alternative lending tool assessing creditworthiness using alternative data in low- and middle-income countries and has been in operation/existence for at least the past 12 months ● Speaks English ● Informed consent obtained ● Over 18 years old
Exclusion criteria	● The fintech has been in operation/existence for less than 12 months ● Does not speak English ● Informed consent is not obtained ● Under 18 years old

I utilized purposive sampling to identify and recruit the study sample, which followed a landscape review of ML-based alternative lending companies to identify companies to reach out to. To recruit, I sent emails or messages through LinkedIn with the interview request. I also attended events whereby target interviewees were speaking or where their attendance was likely based on conference or workshop topics. I also utilized snowball sampling, whereby several participants introduced me to others who meet the inclusion criteria. Recruitment was tedious and challenging. Requests for interviews, particularly to people whom I did not know, were largely not acknowledged. In sum, over 160 people were contacted with an interview request to obtain the 25 interviews (see Table 4 for an overview of interview participants).

Table 4. Interview participants overview

Category & description		Participant count	% of participants
Interviewee demographics[5]	Gender: Male participants	18	72%
	Gender: Female participants[6]	7	28%
	Location: US-based	15	60%
	Location: European-based	2	8%
	Location: India-based	5	20%
	Location: Kenya-based	1	4%
	Location: Other	2	8%
Role of interviewee	Fintech employee (leadership)	12	48%
	Fintech employee (data scientist)	2	8%
	Fintech employee (leadership & data science role)	5	20%
	Investor or funder	4	16%
	Partner researcher[7]	2	8%

Interview participants were largely male (72%). Of the seven female participants (28% of my sample), one was in a leadership data science role, and three were in leadership (non-data science) positions (see Table 5). The remaining three were funders, investors or research partners. These numbers are reflective of the fintech industry more broadly and the lack of female representation in the industry, particularly in leadership and data science roles. Fintech is one of the few sectors that combines two historically male dominated industries: finance and technology (FT Partners, 2022). Globally, only 1.5% of fintechs are solely founded by women (Findexable, 2021). Of the fintechs interviewed, only one fintech was solely

[5] I do not offer race, ethnicity, or nationality of the interviewees because this was not explicitly asked.
[6] Of the 7 female participants, 1 was in a leadership and data science role, and 3 were in leadership (non-data science) positions. The remaining 3 were funders, investors or research partners. This reflects the lack of female representation in the industry more broadly particularly in data science roles.
[7] The two partner researchers interviewed also worked at umbrella organizations that invested in fintechs in the ML-based credit assessment space. Their roles, however, were research related.

founded by a woman (although I did not interview her, but rather another female leader at the company). Meanwhile, of approximately 1,000 of the best-funded fintechs globally, 89% of board seats are filled by men, 81% of C-level roles are filled by men, and women's roles in fintech tend to be in areas of HR and marketing (Findexable, 2021). Interestingly, of those fintechs that are seeking to drive financial inclusion, there is greater gender representation in leadership: A global innovation competition (Inclusive Fintech 50) hosted by the Center for Financial Inclusion highlighted 50 "emerging inclusive fintechs with the potential to drive financial inclusion", of which slightly over a quarter of applicants were women (*Inclusive Fintech 50*, n.d.). Overall, my sample reflects the lack of gender diversity in the fintech industry, with slightly higher numbers than overall fintech diversity averages. This may be linked to the fact that these fintechs are seeking to enhance financial inclusion, a subset of the fintech industry where there is greater gender representation in leadership.

Table 5. Interview participants, by gender

Percent of interviewees in each role, by gender	**Participant count**	**% of women, by role**
Fintech employee (leadership) – Female	3 (of 12)	25%
Fintech employee (data scientist)	0 (of 2)	0%
Fintech employee (leadership & data science role)	1 (of 5)	20%
Investor or funder	1 (of 4)	25%
Partner researcher	2 (of 2)	100%

Interview participants included fintech leaders and data scientists, as well as investors and research partners. The majority of interviewees were in leadership and management positions outside of the data science team (48%), followed by people occupying leadership roles in the data science arm of the organization (20%), followed by investors (16%) and then other data scientist team members (8%) and partner researchers (8%). The vast majority of those interviewed are US-based, reflecting the dominance of fintechs from the US in the ML-based credit assessment industry in LMICs. The US – and particularly the San Francisco Bay Area – has a vibrant startup and fintech ecosystem that facilitates access to venture capital and provides proximity to technology innovation (O'Reilly et al., 2025). Indeed, the US and Canada remains the region with the largest number of fintechs representing 38% of the global total and has accounted for nearly 40% of all fintech funding linked to its larger investor base which plays a critical role in nurturing fintech ecosystems and helping fintechs scale (O'Reilly et al., 2025). Of the 40% of interviewees who are not based in the US, the majority are American citizens, referenced work history in the US, currently work for US-based fintechs, and/or have connections to the US, such as through education histories.

Many of the interviewees, particularly those in leadership roles, had backgrounds in banking and/or microfinance. For example, various interviewees mentioned banking and investment roles as organizations spanning Credit Suisse, Citibank, and Morgan Stanley; as well as digital finance companies such as PayPal. One interviewee also shared their role in helping to start the first credit bureau in India, a spinout of TransUnion (from the US). He later worked for another bureau that branched to India from the US, Experian.

Upon agreement for an interview, informed consent was obtained with anonymity and confidentiality of interviewees maintained through removal of names and other unnecessary personal details from interview transcripts, while files were password protected and encrypted. The interviews took place from August 2023 through January 2024. All interviews (except for one) took place virtually, using video conferencing software, and were transcribed using a speech-to-text transcription service. One interviewee did not consent to their interview being recorded or to quotes being used.

The interviews were organized around a Topic Guide that included a set of broad topics and a set of questions for discussion organized around key concepts (see Appendix 1 for the Interview Questions and Topic Guide). The topics delved into conceptualization, design, and management of the AI tools. More specifically, the interviews delved into how the AI tools were conceptualized, values embedded, and approaches around fairness the tools are optimized for; AI model design and management choices; and implications related to gender. Following each interview, I wrote a one-page memo that includes: (A) important moments from the interview (including ethnographic detail); (B) an analytical section outlining how the interview fits with previous interviews (including where it concurs and diverges); and (C) a methodological section noting the perception of the relationship to the interviewee allowing me to reflect on my positionality and integrate reflexivity.

The full corpus of material analyzed included the interview transcripts and analytic memos developed post-interview. The interview transcription was done using a (post)positive perspective focusing on manifest content of the interview including *what* is said by transcribing intelligent verbatim (removing verbal and non-verbal details that don't add additional meaning). I conducted thematic analysis, whereby I developed, analyzed, and interpreted patterns across the data by coding it to develop themes. I adopted an abductive approach by combining deductive components as well as inductive using a grounded theory method to allow important themes to emerge. For example, I initially created several codes based on my theoretical framework and literature review (e.g., "gender in teams", "defining bias"). As new themes emerged from the data (e.g., "women as better repayers"), I added new codes ensuring both theory-driven and data-driven themes were captured. I drew on reflexive thematic analysis (Braun & Clarke, 2006), whereby I generated initial codes after assessing all interview data and subsequently reviewed and refined themes to identify patterns and core elements in the data. This iterative process informed the development of my codebook and analytical writing. I ensured my coding frame was coherent across the dataset and captured key features for my research focus. Also, I coded terms to categories of substantive (specific details relevant to analysis), theoretical (causal accounts, unintended consequences, turning points) and methodological (e.g., good quotes) (Gerson, 2020). I used Nvivo data analysis software to support the process.

4.3.2. RQ2: Survey and computational social science

My second research question asks: How are benefits and challenges related to accessing and using

ML-based alternative lending tools experienced similarly or differently for men and women? The analysis draws on a mixed dataset comprising 342 user surveys and 234,740 Google Play reviews, both from users in Kenya of the same app. While not triangulation in the strictest sense, combining self-reported survey data with unsolicited user-generated content allows for thematic comparison and broader insight. Kenya was selected as the case for empirical and strategic reasons: it was the first major market for such technologies, remains one of the most saturated, and exhibits persistent gender gaps in digital literacy and mobile internet adoption (GSMA, 2024). In addition, I was able to partner with a fintech company for the survey, whose largest market is Kenya.

The survey was implemented in collaboration with my fintech partner (see Survey Questionnaire in Appendix 2). A full description of recruitment procedures, sampling targets, and ethical protocols (including informed consent, compensation, and data confidentiality) is provided in Chapter 7. Ethical clearance for the study was obtained from University of Oxford, and care was taken to ensure participant anonymity and voluntary participation.

I complemented the survey analysis with computational social science methods, specifically employing natural language processing (NLP) and text analysis of a large user review dataset scraped from Google Play. The particular fintech[8] (i.e. my research partner) had over 1.8 million reviews as of August 2024. Android devices dominate the market in Africa and Asia (*Mobile Operating System Market Share Africa*, 2024; *Mobile Operating System Market Share Asia*, 2024) and Google Play is where these apps are located for download, hence I am using Google Play. I scraped all reviews that were available in January 2024. The dataset includes reviews provided on a date ranging from April 2020 to August 2024. In total, 234,740 reviews[9] were scraped; of which, 38,344 had names associated with the review and I was able to assign gender in 18,799 reviews. Details of the computational methods and models are outlined in Chapter 7.

[8] The fintech is anonymous per a Non-Disclosure Agreement signed as part of the process to conduct a survey in collaboration with the fintech.
[9] This number reflects the reviews that included text.

I employed a multi-method analytical strategy grounded in my conceptual framework drawing on gender and development research, feminist economics, and technofeminism. The analytical strategy included descriptive statistics, regression modeling, and mediation analysis to test whether women reported greater challenges or benefits and whether these were shaped by factors such as financial access and digital literacy. I then employed computational social science methods, specifically text analysis, on the large user review dataset, which offered unstructured data for complementary insights related to my hypotheses. This text analysis combined both exploratory and comparative components: I identified topics that emerged organically from the unstructured user reviews, and also examined whether and how topics that surfaced in the survey findings were reflected in the review data. This approach enabled me to generate new insights from the reviews while also assessing the extent to which patterns identified in the survey were echoed in unsolicited user accounts. I also conducted text analysis on open-ended survey responses to contextualize and deepen interpretation of the structured survey results. Finally, alongside the hypothesis-driven analyses, I include a set of exploratory findings from the survey data that provide context on app users and their loan behaviors.

4.4. My Positionality and Ethical Considerations

4.4.1 Positionality

I reflect on my own positionality to understand my identity, values and life experiences that enter into the research process. I am a white woman living in Berkeley, California, who is a doctoral student at University of Oxford studying impacts of algorithmic-facilitated lending, including gender bias, in LMICs. It was my own identity as a white American-Canadian woman based in California that contributed to my initial interest in my research questions. As someone who had studied algorithmic bias prior to beginning my dissertation, I was frustrated by the lack of research on algorithmic bias outside of the US and Europe, despite the proliferation of AI globally, as well as a lack of research on gender and AI in particular. AI tools are often developed or managed in the West and implemented in the Global South under labels of being “for good” without being analyzed for their implications and potential unintended consequences. My identity also informed my research approach of wanting to study fintechs themselves, many of which are based in my local Bay Area region or have roots and connections to Silicon Valley, while also wanting to ensure I gathered information from the voices of people using the tools – including women – themselves.

In conducting research related to international development, particularly as it relates to question 2 which examines the voices of users, I am mindful of my whiteness and nationality. As a white woman who lives in and grew up in the US, in international development spaces, I am intricately aware of values I hold around feminism, justice, and gender equity. I am mindful of histories of well-meaning “Western feminists” that can import Western ideals in different cultures and contexts, represent women in the Global South as “victimized”, and/or posit the “third world woman” as a monolithic category for analysis (Gandhi, 2019; Mohanty, 1984). I am also mindful that colonial ideology has, in some contexts, represented new forms of freedom for women (R. J. C. Young, 2020). In studying “AI for Good” and employing feminist and postcolonial theory, I hold these histories and understandings in mind. I recognize that my own lived experience as an American woman shapes my perceptions of feminism, as well as values of egalitarianism and liberal feminism. In my theoretical basis, I am careful to draw on theorists from the majority world. I draw in intersectionality to ensure that I am not presenting women as a monolith in my assessments of gender impacts and bias. I am careful in my analysis to assess my own positionality frequently, question myself, and question my presumptions. I am also mindful not to attach normative prescriptions to findings. For example, even if ML-based credit assessment tools increase access to finance overall but not in gender equitable ways, women still can appreciate access to them for themselves or their families and communities. It is not my place to say whether they are ultimately “good” or “bad”, nor would I seek to paste such a label. Rather there are tradeoffs. In my research and ensuing recommendations, I seek to lay out those tradeoffs while encouraging prioritization of local voices and perspectives in considering tradeoffs.

4.4.1.1 Positionality related to RQ1 and qualitative data collection

My identity and lived experience played a critical part in gaining access to the elite interviews conducted. My background in international development, including having worked in international development on financial inclusion and digital inclusion in the past (including with tech startups), gave me an understanding of opportunities and challenges fintechs can face in the international development arena. While conducting the interviews I also held a position as the Gender & AI Fellow at the US Agency for International Development (USAID). My history in international development and current roles meant that I had connections I could leverage to get introductions and access to spaces where prospective interviews would be. As I connected with prospective interviewees most often through LinkedIn, my current academic and professional positions certainly supported me in interviewees agreeing to participate. This was also very helpful in getting a fintech company to agree to partner with me on the survey. My background and ability to understand some of the constraints and opportunities of fintechs in this space helped me to engender trust with the participants and partner fintech. Being based in the Bay Area was also helpful to meet some prospective interviewees in person at different local events and workshops. In particular, one interviewee who is the CEO of one of the largest fintechs in this space, agreed to an interview after I introduced myself at a panel event at the Haas School of Business at the University of California, Berkeley (where I also worked during the period of this research). Prior, I had been unsuccessfully attempting to gain an interview from the Co-Founder for weeks.

Overall, despite my background and success in reaching 25 interview participants, gaining access to interview participants was very challenging and included countless hours of LinkedIn requests and messages, attending events and workshops that prospective interview participants may attend, and leveraging mutual connections for introductions as much as possible. Several interviewees looped in representatives from legal prior to participating or backed out last minute when they did not get approval to speak with me. I included interviewees who partner with fintechs, including investors, funders, and partners, to expand the pool of potential participants and gain a more comprehensive understanding of my research questions.

In conducting the interviews, my positionality also influenced what interviewees told me and the type of information I was able to collect. I was worried that interviewees may not be fully honest with me in regards to challenges they faced, particularly linked to gender and any data around their customers. I thought my association with a funder, USAID, could impact what they told me, while I know there can be hesitation generally around sensitive topics like gender differences in lending that could indicate bias. Some interviewees did seem to follow more of a corporate script, while others were more open and honest about challenges and their feelings or perceptions. Interestingly, I found that one of the most honest and insightful interviews was with the CEO of one of the largest fintechs in this space (who also requested not to be recorded or quoted). This honesty may have been because I was a doctoral student, as opposed to a faculty member, so they felt they could be more honest with me. This same interviewee also expressed negative feelings about journalists who they felt may pursue “gotcha” stories and present skewed information in order to sensationalize a story and grab headlines. This interviewee, since we met at UC Berkeley, knew of my association to Berkeley and expressed frustration at one point in the interview when discussing fairness (making connections to “social justice warriors”). Similarly, another male interviewee who lived in the US but was from Europe, when I asked about fairness, laughed and noted it was a “US kind of question”. This highlights how there was some perception of me as a young American woman who cared about social justice, linked to the topic of my research and – if they did background research on me or looked at my LinkedIn profile – work I have done in the past. Of course, in my research my goal is to be as objective and open as possible, while being interested in finding different types of results that make me question my own hypotheses and assumptions. Several times interviewees asked for something to be off the record highlighting some of the sensitivities of what we discussed.

In the interviews, I presented myself as an eager doctoral student keen to learn more about the topic area (which is accurate), while also being clear about how data from interviews is anonymized and aggregated for analysis. I worked to make the interviewee comfortable by being friendly and making small talk to start. I feel that I was successful in this effort. Many of the male interviewees were older, and while I generally felt treated as a peer, for some men, particularly older Indian men, I became very aware of my gender and age. Several male Indian interviewees said for me to contact them if I ever came to India and one invited me to India to visit. Overall, I think being a young, white, American female served me in being able to get access to interviewees who were more honest with me as they did not see me as threatening.

4.4.1.2. Positionality related to RQ2 and fintech partnership

My positionality informed my process and approaches for collection of quantitative data. Using technical skills in Python, I was able to scrape the large review dataset. In regards to securing survey data, I was able to successfully secure a partnership with one of the largest fintechs in this space. My identity as an American doctoral student in international development with a history and focus on gender, while also being experienced in conducting research internationally such as in a prior position at the International Center for Research on Women (ICRW), was helpful. The fintech company had recently made a strategic decision to focus on women as a target customer segment, so the timing was good to approach them for a partnership. I worked with a woman at the company in developing the partnership agreement and the lawyers of the company were looped in for the development of the partnership agreement and NDA. In addition to the survey, I asked them if they could provide the Google Play review data by gender (as I told them I scraped the ~235,000 user reviews of the app in Kenya and was using a name-to-gender service in Python to assign gender). However, they also did not know the gender of the users who left reviews. Overall, it was a positive experience of mutual respect and interest.

My positionality informed the survey design and its iteration. I designed and completed the survey questionnaire draft after collecting interview data, which helped me refine my survey questions. I drew from my prior research experiences conducting surveys with startups on topics of women's empowerment and technology while working at ICRW. For example, at ICRW I designed and implemented surveys that leveraged different scales to assess impacts, particularly as it relates to technology, digital inclusion, and women's empowerment. I was able to conduct a pilot of the survey with target users and staff of the fintech given my good relationship with the fintech partner. Piloting the survey was very helpful in understanding where the survey was confusing and how to clarify my questions. Despite my lived experience being very different from target respondents of my survey, piloting the survey enabled me to refine and finalize the survey in ways that were clear to my target respondents.

Finally, my positionality played a role in interpreting the results of topic modeling conducted for both the open-ended survey responses and review data. Topic modeling identifies patterns of word usage and clusters them into a specified number of topic sets, requiring researcher interpretation to determine overarching themes. My interpretation of those themes was informed by insights I gleaned previously from the interviews, other quantitative survey data, and my theoretical framework. My interpretation was also shaped by my knowledge of gender differences in digital inclusion, digital literacy, and financial literacy. As someone with experience in gender and technology research, I was particularly attuned to how gendered patterns in my data may reflect broader societal dynamics, such as gender disparities in access to smartphones or financial literacy.

4.4.2. Ethical considerations and limitations

I faced several ethical questions and considerations. All interviewees with the exception of one were comfortable being recorded and quoted. One interviewee declined to be recorded and quoted. I assured interviewees that attribution would be anonymized and recordings are used only for research purposes that are deleted after transcription. Given that algorithmic-facilitated lending in LMICs, while a growing space, has clear industry leaders, I keep my quotes anonymized to not expose particular organizations and research participants. For the fintech company that I partnered with to do the survey, I agreed to allow them to choose whether they would like to be anonymous following their review of the findings. They were also aware that I had scraped the Google Play review data. This allowed them to be comfortable in partnering with me and is a process they have followed similarly with other researchers with whom they have partnered. Partnering with them was very helpful in both gaining access to users, while also being able to ask follow up questions. I found the partnership to be complementary, as they were authentically interested in my findings. When sharing findings with the fintech partner, I include recommendations based on the results to support their expressed interest in ongoing improvement.

There is always a power imbalance between the researcher and interviewees. Particularly when studying marginalized communities, the power imbalance tends to fall towards the researcher. However, in the context of elite interviewing and "studying up", the power balance is flipped. I maintained a friendly and curious demeanor, while also being mindful to push as appropriate on certain areas. In some cases, interviewees would show up late or not at all, and as I was often fitting into a busy day of theirs, I was careful to keep to time or ask politely for slightly more time if needed.

In conducting research related to gender, it is critical to acknowledge that gender is not binary. Rather, gender is a spectrum and refers to social constructs of norms, behaviors, and roles that vary within and across societies and time. Gender includes non-binary and gender non-conforming individuals. I purposefully kept interview questions open around gender to allow for respondents to discuss considerations related to nonbinary people. In questions asking about demographics of users, I noted "M/W/nonbinary" as categories for which I asked for information. I also included nonbinary gender identification in my survey. No interviewees referenced nonbinary people, nor did I get any information on nonbinary people in my survey data. This is fine for my research as I explore the ways that gender norms and biases come into play in technology design, development, and management. Future research may target and further explore how nonbinary people experience ML-based alternative lending systems.

There are several ethical considerations as it relates to the Google Play review data and NLP analysis. First, scraping data online, even though the data is publicly available, has ethical considerations. Scraping online data has been a valuable tool for academic research across multiple disciplines, with the Internet being a rich source of data and information about humans (Rennie et al., 2020). However, scraping information from online has ethical questions. For example, data scraped that reveals people's locations raises ethical concerns around tracking users (Hsu, 2018). Privacy standards are not always clear. Relatedly, people who leave online data did not necessarily grant permission for their opinions to be used for research purposes despite the data being publicly available. These are tensions many data scientists and researchers face given common practices of web scraping (Rennie et al., 2020). In the context of Google Play review data, comments do not reveal sensitive information and location data is not collected. I also deleted names, where names were attached to comments, after assigning gender.

The NLP analysis relies on assigning a gender label based on one's name to app reviews[10], which has ethical implications. First, it relies on gender as binary; but gender is a spectrum. Secondly, it also assigns someone an identity label which may or may not be their identity. While these tensions remain and there can be inaccuracy regarding assignment of gender, academic research on the misclassification rates of name-to-gender inference services (particularly Python's gender-guesser) reveal that across geographies and cultures misclassifications are less than 2% (Santamaría & Mihaljević, 2018). This misclassification can be higher in non-Western countries. This minor bias is not problematic for the research as utilizing big data allows systemic patterns to emerge that can wash out smaller effects, while I am also assessing general trends related to how women and men access and use the app. I mitigated this tension through reviewing a random subsample for accuracy by checking Facebook accounts and by being clear about the limitations and explicit about how gender categories are being assigned (Larson, 2017). Lastly, I addressed this tension by interpreting the review data through the primary lens of the survey findings, which included respondents' self-identified gender.

The research also faces several limitations. First, interview and survey responses were self-reported, with research participants self-selecting to participate. For the interviews, I did not observe respondents in the

[10] Note: names included in reviews are linked to one's Google account and devices; thereby tend to include an actual first and last name.

workplace. However, this also allowed me to get a sample across multiple organizations. Second, for the interviews, I used snowball sampling as a recruitment method, which may have led to respondents who already had a heightened interest and awareness of topics related to gender and AI. Future research could dive deeper within a small number of fintechs and incorporate observational data, such as ethnography, so as not to rely on self-reported data. Ultimately, I chose to prioritize a breadth as opposed to a depth of sample by interviewing individuals across multiple fintechs. Limitations and future research are further discussed in Chapter 8.

4.5. Conclusion

In this chapter, I explained how I undertook my research assessing algorithmic-facilitated lending and gender in LMICs. I applied mixed methods, using both qualitative and quantitative methods – including semi-structured interviews, a survey, and big data analysis – to answer my research questions. The research questions and methods are theoretically grounded drawing on feminist theory, postcolonial theory, and STS. My overall research approach allowed me to dissect the ways that ML-based alternative lending tools by fintechs reinforce or mitigate gender inequitable access to finance in LMICs, despite these tools being closed to outside researchers and people broadly. My research approach equipped me to build a holistic and nuanced understanding of how ML-based alternative lending tools can impact financial inclusion and gender implications, including illuminating the "ground truth" around root causes of how the AI models are assembled and examine power structures therein.

CHAPTER 5: MINDSETS AND MANAGEMENT – CAN WE LEAVE IT TO THE MACHINE?

5.1. Introduction

In 2012, John[11] became obsessed with a digital lending side project. His goal, as shared when telling me his story in an interview, was to help unbanked entrepreneurs access finance by using machine learning to assess alternative data on peoples' smartphones – which were proliferating rapidly in LMICs at the time – and provide a credit score. As an American man living and working in rural Kenya in his mid-20s, he volunteered at a microfinance charity and recognized the transformative power of capital for micro-entrepreneurs, many of whom remained stuck in poverty. He also witnessed the barriers that kept poor, trustworthy borrowers outside of formal financial systems, despite their hard work. This realization inspired his career trajectory and mission to unlock capital for those excluded from financial opportunities. Since launching his business nearly a decade ago, it has reached millions of people in Kenya and beyond. However, as this and other similar algorithmic lending innovations scale in LMICs, it raises critical questions about how the use of ML in credit assessment can enhance financial inclusion and reinforce or mitigate gender biases in access to finance.

This chapter examines the underlying logics, design choices, and management decisions of ML-based credit assessment tools by fintechs to explore if and how they embed or challenge gender biases, and the impact these systems have on gender equity in access to finance (RQ1). This chapter is informed by qualitative interview data with corporate leaders, investors/funders and data scientists at fintechs developing and managing ML-based alternative lending apps. By investigating these dynamics, this chapter seeks to contribute to ongoing debates about the role of AI in shaping inclusive financial systems,

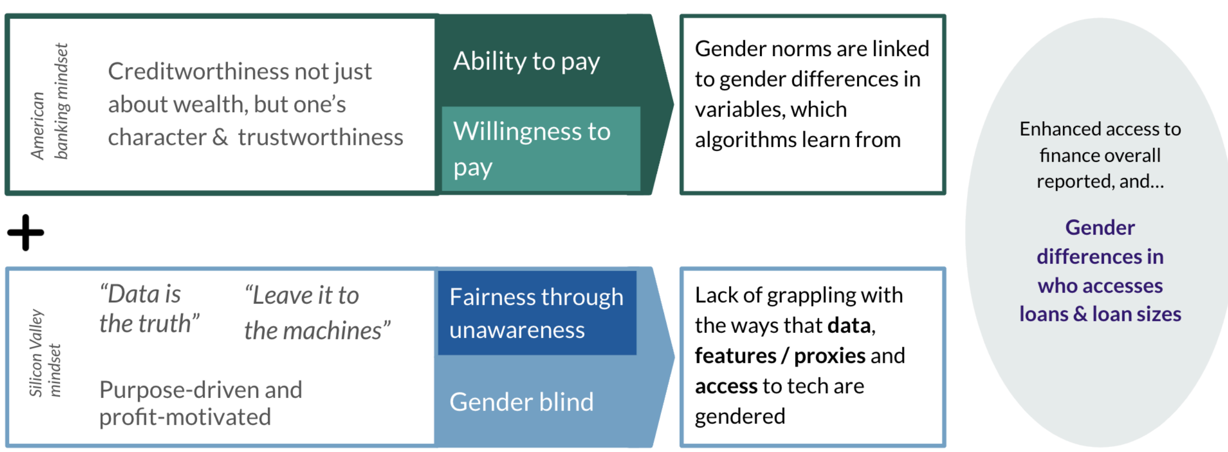
[11] All names of interviews are pseudonyms.

the moral choices that come up when introducing algorithmic tools into daily life, and the relationship between algorithmic decision-making and equity more broadly.

In this chapter, I draw on my theoretical framework, which integrates insights from STS, feminist theory, and postcolonial theory that emphasize how technology development is not neutral, but rather shaped by individual choice and power dynamics (Kuhn, 1970; MacKenzie & Wajcman, 1999; West, 2019). Specifically, I use Donna Haraway's *situated knowledges* and Ruha Benjamin's concept of *default discrimination* to analyze how ML-based credit assessments in LMICs embed or challenge gendered power structures. Haraway's situated knowledges posits that all knowledge is produced from specific, socially, and historically situated standpoints (Haraway, 1988). I apply this to critically assess how the design of ML-based credit assessment tools reflect the perspectives and priorities of their creators, which are shaped by the cultural and institutional logics of American banking and Silicon Valley. I explore how creditworthiness, as operationalized in algorithmic tools, is not an objective measure, but a product of historical and cultural standpoints. Technofeminism supports this analysis by exploring how the social context of technology developers influences the design choices embedded in these tools, enabling me to critique and unpack choices by algorithmic developers and managers (Wajcman, 2006). Benjamin's concept of default discrimination informs the gap between intention and outcome in technology design. It critiques how even well-intentioned algorithms, designed with the aim of inclusion, can perpetuate existing biases if they fail to address the underlying structural inequalities embedded in the data and development processes (Benjamin, 2019a). I use this to explore how "'gender blind" approaches in ML-based credit assessment can – largely unknowingly – lead to the perpetuation of gender inequities and mask reproduction of existing power hierarchies. Together, these frameworks and concepts allow me to unpack the underlying logics, design choices, and management decisions of ML-based alternative lending tools by fintechs, and critically assess if and how they embed or challenge gender biases and influence gender equitable access to finance.

After briefly describing my methods (which are further fleshed out in the Methodology chapter), this chapter is organized into two sections outlining findings. The first section examines the underlying American banking logics that inform the definition of "creditworthiness" and how this translates into algorithmic choices, alongside Silicon Valley logics that promote data-hungry approaches and affirm beliefs that ML is objective. In the second section, I explore how these logics inform design choices related to building gender "blind" algorithms that lead to a lack of grappling with the ways data, features or proxies for creditworthiness, and access to apps are gendered and reflect inequities. I outline fintechs' perceptions regarding why women receive less loans and at lower amounts than men, before providing socio-structural explanations to expose how gendered inequities are reproduced within fintech technologies and through lending decision-making. I conclude with a discussion on my findings, including highlighting an *objective algorithm paradox*, in which the belief in ML's objectivity results in being "blind" to certain demographics in ways that hide the gendered nature of algorithms and perpetuate discrimination. I expand on existing theories of algorithmic bias and gender equity by introducing the concept of *encoded gender norms* and exploring its relevance in the context of ML-based credit assessment and algorithmic decision-making. Figure 1 outlines the findings and provides an illustrative overview of the chapter.

Figure 1. Overview of findings



5.2. Methods

The findings in this chapter are drawn from 25 semi-structured interviews conducted with corporate leaders, data scientists, investors, and research partners at fintechs developing and managing ML-based alternative lending apps. The questions explored key topics, including the conceptualization, design, and management of AI tools, and focused on underlying logics, mindsets and values. Questions examined AI model design and management, including overall implications and lessons related to gender. I utilized purposive sampling to identify and recruit the study sample, which followed a landscape review of ML-based credit assessment fintechs to identify companies to reach out to. To recruit, I sent emails or messages through LinkedIn with the interview request. I attended events whereby target interviewees were speaking or where their attendance was likely based on conference or workshop topics. I also utilized snowball sampling, whereby several participants introduced me to others who meet the inclusion criteria. Ultimately, over 160 participants were invited for interviews and 25 interviews were conducted between August 2023 and February 2024 (see demographics of interviews in Table 4 of the Methodology chapter). Each interview lasted approximately one hour, allowing for in-depth discussions with participants. All interviews (except for one) took place virtually, using video conferencing software, and were transcribed using a speech-to-text transcription service. One interviewee did not consent to their interview being recorded or to quotes being used.

In analyzing the data, I utilized an abductive approach through combining deductive components building from my theoretical basis and analytical framework, as well as inductive using a grounded theory approach to allow important themes to emerge. NVivo software facilitated the coding process. Drawing on reflexive thematic analysis (Braun & Clarke, 2006), after familiarization with the data round, I developed an initial generation of codes. I developed and reviewed themes from the codes before subsequently refining my codebook and writing. To ensure reflexivity and positionality, I consistently reflected on how my own identity shapes my knowledge and assumptions in the interviews, including through incorporating reflection into post-interview memos.

5.3. Findings (A): Underlying logics and mindsets

Using ML to assess the credit of unbanked and underbanked people in LMICs emerged in the last 15 years. Several of the leading and largest fintechs in this space have roots in the Bay Area and Silicon Valley. My findings reveal two primary mindsets and logics rooted in American banking and Silicon Valley that guide the design and management of ML-based credit assessment tools. First, these alternative lending tools incorporate concepts of creditworthiness that stem from American history and values, wherein one's creditworthiness is not simply about wealth, but also one's character and trustworthiness. Simultaneously, interviewees hold Silicon Valley logics, in which more data is better, data is the "truth", and leaving it to the machines is best given their perceived objectivity.

5.3.1. American banking mindsets

5.3.1.1. The purpose of ML-based credit assessment

Fintechs are grappling with the same question that has been at the forefront for bankers for centuries: will this person repay my loan? As Kamal, an interviewee based in India who runs an Indian fintech company, put it: "ever since money was invented, lending was invented. It's not a new business, right? … But how do you give a loan? How do you evaluate the person?" According to interviewees, the primary goal of these tools is to accurately evaluate whether a person is "creditworthy" to receive a loan, and if so, under what repayment terms. Creditworthiness, as operationalized in these systems, draws on two components rooted in American credit evaluation traditions: (1) the ability to repay and (2) the willingness or intent to repay. Kamal continued, "[There is] sort of this universal law… [For] every bank in the world, every lender in the world, fundamentally there are two things… The ability to repay and the intent to repay." Interviews revealed that, across fintechs, ML tools are designed to predict both aspects in order to generate a credit score (though not typically disclosed to app users), which ultimately informs loan approval decisions and repayment conditions. ML-based credit assessment tools are thereby assessing one's creditworthiness tied to the logic that people, even those who are poor, can be creditworthy – as creditworthiness is tied to one's character and willingness to repay as opposed to simply their level of wealth. This echoes the dominant concept and notion of creditworthiness in banking today that has roots in America, and was born under pursuit and promise of meritocracy, in which anyone can reach the highest levels through their own ability and skill, as opposed to external factors such as the social class of one's parents.

While this conception of creditworthiness has roots in America, it is not limited to US soil. Rather, the concept was and continues to be exported through credit scoring and credit bureaus. Amir, based in India, discussed his history working in the first credit bureau in India – which was a spinout from the American credit bureau, TransUnion – and later at another American credit bureau spinout in India, Experian. He said, "After working with the TransUnion Civil, I worked with Experian, which again in [the] US is one of the largest [credit] bureaus… I worked again on setting up Experian's business in India… When I was in the Bureau, I realized there is huge potential in India to [address the unserved and underserved]." Amir's story illustrates how American banking logics embedded in ML-based credit assessment tools are not delegated solely to Americans or US-based firms.

There are several ways in which ability to pay and willingness to pay are examined and encoded in the algorithms by fintechs.

5.3.1.2. Operationalizing creditworthiness: Ability and willingness to repay

Ability to pay is about whether one has the financial means to repay a loan. Ability to pay is largely an assessment of one's job, income, cash flow, and broader financial situation. Kamal noted that estimating ability to repay is, "fairly simple. How much money is [the person] generally making every month? How much is she spending every month? And how much do you have in the bank?" Within this, having an understanding of the stability of one's economic and job situation is important. Fintechs may consider whether one has an informal or formal job, as well as a cyclical/seasonal job or steady job, the latter in both cases being more favorable. Fintechs rely on alternative data gathered through smartphones and utilize ML to estimate aspects of one's economic and job situation, but may incorporate questionnaires and ask for supporting documentation (e.g., pay slips). Cash flow can be assessed based on how much one is making and spending monthly, which can be tracked over a period of time through assessing one's transactions, withdrawals, and deposits (whether in formal bank accounts, other more informal digital accounts such as MPesa, or through text messages documenting different cash flow metrics). Fintechs may leverage ML to estimate cash flows and job stability through a variety of data and variables. Amir, the co-founder and leader of a fintech in India, explained how they identify income and cash flow:

> "Some examples of the variables [to arrive at income estimation], could be rent payment, could be the location where the person is staying… depending on the profile's data we are able to compute the income. The biggest point here is, how do you look at various transaction history to build a cash flow? Transaction history could be a banking transaction, it could be [an] insurance transaction, it could be a telecom transaction, it could be a rent transaction, it could be bank withdrawal… Then I can actually tell you at what point in time you had what balance… Then I exactly know whether the income is cyclic or regular…. You know exactly how much of [their] income is coming, how much spend[ing] is there."

Based on the various information and variables, one's ability to pay is assessed. A stable and positive cash flow alongside consistent income would be considered good for the algorithm that is assessing one's ability to pay.

Willingness – or intent – to repay is different. At a high level, this assessment is about whether a person is honest and trustworthy. Willingness to pay could come from various sources, and includes an assessment of one's behavior, particularly one's financial behavior and whether a person is financially disciplined. In addition to financial behavior, it also includes an assessment of the person's reason for taking out a loan (e.g., for education, paying back another loan) and whether one intends to repay or is a fraudster. Assessing willingness to pay can be considered more challenging. Kamal asked: "Are they generally an honest person? ... That's a very tricky question because now you're questioning someone's morality, right?" Willingness to repay is applying numeric understandings to questions of one's honesty, trustworthy, and morality.

There are several common approaches in assessing willingness to repay. This can include assessing one's behaviors (e.g., whether one takes out many loans based on how many financial apps they have on their phone, whether one is a gambler based on their phone apps), using psychometric testing to understand one's psychology, and assessing one's relationships with others. If credit scores and credit histories are available on a person in a certain country location, those may be combined with one's smartphone data. In many cases and depending on the country, credit scores are not available.

Psychometric tests are meant to identify an individual's personality, mental capabilities, and behaviors. As willingness to pay is about a person's trustworthiness, several interviewees note how assessing personality is thereby important. This might be done through a questionnaire and/or through an assessment of one's psychology using data acquired via their phone. Two interviewees discussed the approach of psychometric testing in assessing willingness to pay. Rakesh shared his company's approach of integrating a questionnaire and using ML "to come up with an integrity coefficient based on all the responses… It predicts… this level of integrity, and there's a very high correlation between integrity and willingness to pay alone." Chris, who runs a fintech that provides ML-based credit assessment tools to various organizations globally, discussed how they assess a person's personality using smartphone data and a gamified app. They "pull data from all sources that we've learned contribute to defining and describing someone's personality… We are measuring intelligence, reliability, stability, consistency, integrity. We've got like 20 of them… The algorithm picks [up] on [information] and… it tells us [what] correlates to those features." In short, one's behavior and trustworthiness is assessed in a variety of different ways with alternative data and examining correlations. While prior studies have examined the ways that psychometric testing can support credit scoring in LMICs (Arraiz et al., 2017), there is a lack of research exploring the bias and ethical issues that can be present in psychological assessments for credit scoring.

5.3.2. Silicon Valley Mindsets

Two core beliefs associated with Silicon Valley are reflected by fintechs and present in ML-based credit assessment tools: the assumption that machines are objective and the idea that data is a reflection of truth. These two beliefs, which go hand-in-hand, are pervasive amongst fintechs in the ML-based credit assessment space and were reflected across the interviewees, including amongst those geographically further from Silicon Valley (e.g., in Kenya and India). Other prevalent Silicon Valley ideologies that shape these tools in LMICs include an emphasis on innovation and being purpose-driven.

Similar to American banking logics, Silicon Valley logics are not restricted to that physical geography. This can be partially linked to the prominence of Silicon Valley globally and its importance as an entrepreneurship and technology epicenter, from which innovations and ideas flow and are exported. Even if fintechs are not based in Silicon Valley, several leaders and investors share roots or connections to the Bay Area. For example, Kamal, an Indian fintech leader, went to grad school in the Bay Area of the US where he launched several startups, including one which sold to a major information technology company before he moved back to India where he continued working on a variety of digital ventures before co-founding his fintech. I now explore each of the Silicon Valley beliefs, including how they inform decisions and implications related to gender.

5.3.2.1. "Leave it to the machine"

Across the interviewees, there is a belief that ML tools are objective and less (or no) human involvement or interference is better. This is reflected both in the development of algorithms, as well as trusting in the technology over time. When it comes to developing algorithms for credit assessment, several fintechs discussed disbursing small loans, largely randomly, to see who repays. By distributing loans randomly, fintechs generate real-time data on repayment behavior in specific geographies, rather than relying solely on historical data to predict creditworthiness. This experimental approach allows them to observe what variables correlate with repayment, and these insights are then used to develop and refine the algorithms applied to future applicants. Other fintechs rely on existing or historical data to train their models, using patterns from past borrowers deemed creditworthy to develop algorithms that are then applied to new applicants. In both approaches, ML identifies patterns among individuals deemed creditworthy and uses these patterns to develop algorithms that are applied to new applicants. Dante, who is based in the US and works for a fintech that develops ML-based credit assessment models for use globally, noted their reliance on ML to identify the variables most predictive of creditworthiness, which are then used to build models for evaluating new potential borrowers:

> "Depending on the target that we want to analyze, we'll leave it to the machine to really identify the one variable that has the highest information value, the lowest correlation with other features and also the highest stability… Something we know… If a customer has more than 12 apps in the finance category, they tend to be more than two times riskier than somebody with less than five apps in the finance category. Now, how to explain it from a psychological point of view, [a] behavioral psychology point of view? [It] is kind of challenging because we really leave it to the machines to assess informational value… We know also that a customer that takes a disproportionate amount of selfies tends to be a worse repay[er] than somebody that takes very little selfies… You can understand, perhaps the vanity. We don't try to really explain the features in a human intelligible kind of way, but we just provide data statistics."

This quote illustrates how the machine is used to find correlations, elucidate information about one's creditworthiness, and affirm ideas about what creditworthiness entails – in this case, through connecting creditworthiness to the number of selfies on a smartphone.

The idea of "leaving it to the machine" builds from the recognition that humans can be biased. Machines are seen as trusted entities to reduce, or even fully remove, human bias and subjectivity. Rohan, in the C-suite of a fintech company based in India, elucidated: "We wanted to reduce the subjectivity. Machine learning became one of the more objective methodologies that we could use because it is largely driven by [an] objective set of things… It removes the subjectivity out of this equation." Other interviews echoed that machine learning is objective, and, relatedly, can remove bias. Sarah, based in East Africa who is in a leadership role at a fintech with operations in four country locations, asserted: "The good thing about the machine learning tool is it also removes any bias because there's no human being who's going through the messages looking at what's right, what's wrong. It just follows a particular path or a particular route."

These sentiments, echoed across fintechs, reflect a popular notion that technology and AI is objective, while human beings are the ones that are flawed, subjective, and/or have bias.

Various interviewees discussed the use of machine learning as a method to remove subjectivity and make lending nondiscriminatory. According to Chris, who is based in Europe and leads a fintech that develops ML-based credit assessment models for use globally: "The [founder] wanted to create something that was unambiguous. Didn't discriminate, nondiscriminatory, fair. And so the value of just pulling metadata and letting algorithms go to work to find out common features and correlations makes it agnostic to the individual… The thing I love about it is absolutely there's no prejudice to it… We're just pulling data from a smartphone and then giving… a score." These tools are painted as an important tool for avoiding prejudice and discrimination.

5.3.2.2. Data is the "truth" & "king"

While machine learning is considered objective, data is seen as the "truth" that feeds the model. These two beliefs go hand-in-hand with each other, as the notion of ML being objective relies on data being reflective of fact or truth. Amir, based in India and a co-founder and leader of a fintech company operating in India, remarked, "We don't want to be biased about certain religions, castes, race. We don't ask [about] any of them… We only look at data. And whatever data suggest. I always believe data is the truth. Data will never lie." This concept of data being the truth does not necessarily consider how data can reflect human choices – for example, where was the data collected from? Who collected it? What data informs the dataset and who is represented amongst that data? In reality, there are various decisions that go into the development of data and selection of datasets.

This idea of data being truthful is also central to the belief that these tools are not biased, or are less biased than traditional lending. Data is contrasted with the status quo in access to finance, which has been rife with unfair biases stemming from humans. Sarah put it, "The fact that we eliminated the issue of someone having to show up or fill forms… we don't have to see them. Because there's always this stigma attached to having to walk into a financial institution. Of course, we have our biases. If I look at you, I'll probably assess and say, hmm, this person possibly looks like they can repay. Even before I look at the data. So the advantage became, I don't need to see you to make a decision. I just need the data to speak to us." Data is seen as both holding and speaking the truth.

Data is paramount for fintechs. Rakesh, who is based in the US and founded a company with operations in Thailand and Pakistan said, "Data is king. So anyone who's interested in machine learning or anyone interested in AI, the starting and the endpoint to me is data." ML algorithms learn from data to make predictions and inform decisions, and therefore, data is central to the quality of algorithms and their value. Furthermore, ML technologies are considered to improve in accuracy with the more data that they have. Given this, fintechs are both data-driven and data hungry. Having more data is tied to considerations of being objective and reliable. More data is seen as better to get to the goal of objective, reliable credit assessment. There is not necessarily an end point in regards to how much data is enough.

As data is "king" and more data is better, there is a sense that all data is valuable and insights can be extracted from various types of data. Michael, an investor of fintechs in this industry, explained that data assessed with ML allows an "extracting of fundamental threads". He continued by explaining the data trails that people have, in a Western context, and what it tells you about an individual:

> "The people that go to watch tennis are more likely to buy a Rolex than a Casio, right? So when it comes to financial transactions and the amount of fidelity that we can extract and infer from whether they bank with Wells Fargo or have a credit union or use a cash account, or whether they take money to Western Union or they use different remittance services… How they pay for gas, whether they do prepaid cards... So the consumer behavior profiling that you get out of putting all this into the data layer, you can run the models against it."

Any and all data is relevant to capture the identity and story of a person, which can elucidate insights about their creditworthiness.

In line with being data hungry and the idea that more data is better, fintechs are increasingly looking beyond smartphones to augment smartphone data. Traditionally the approach has been to use data on Android devices and what is able to be gathered through the app. But that is expanding, particularly as

Google Play store (the store where most apps for Androids are downloaded from) have made privacy regulations for apps more strict, partly due to predatory lending apps (Perez, 2023). Several interviewees discussed the approach of getting as much data as possible, including as much data from as many sources, including telecoms, credit bureaus (where available), other government data sources, and/or other third-party data brokers. Some fintechs have tools to help acquire and merge data sources. While not the focus of this dissertation, it is important to highlight the tension between getting as much data as possible and data privacy or consent. Gathering more and more data about an individual can come with a cost of surveillance and privacy.[12] In the Conclusion, I highlight areas for future research including as it relates to privacy and surveillance implications of these technologies.

5.3.2.3. Purpose-driven & profit-motivated: Do good & make money

Many fintechs operating in the credit assessment space in LMICs, including all those interviewed, position themselves as purpose-driven, often emphasizing missions and goals of reaching people who are underbanked and unbanked with access to finance. They argue that they are, after all, using technology to enhance access to finance for people underserved by traditional financial institutions and approaches. This sense of purpose is central to the founding and operation of these companies. Josh, a data scientist at a company headquartered in the US with operations in several country locations said, "This is an awesome opportunity to use devices that people already have to increase their financial agency and autonomy and give them opportunities that they might otherwise not have to get into traditional brick and mortar buildings." Interviewees shared an excitement for the social impact potential of their organizations and the technologies.

[12] Surveillance doesn't necessarily end when a credit offer is provided (or not). Some fintechs discuss the importance of tracking people over time. This tracking is important to not only understand the creditworthiness of the person, but also to build more personalized strategies for ensuring their repayment. While consent mechanisms are built into the apps, and this is emphasized by fintechs, it is an open question of how much people actually understand regarding usage of their data. Would people accept the level of surveillance that is required by the apps to provide a credit score and potential access to a loan?

By enabling access to customers otherwise left out of financial folds, ML opens new market opportunities that can bring business benefits. As put by Jack, who works at an organization that funds and partners with fintechs: “From a lender's perspective, it’s also good for the bottom line. If you can accurately identify people who are creditworthy that you previously were not lending to, that’s a business opportunity.” Lenders are, after all, in the business of making money. At the end of the day, while innovation is prized and purpose is a core driver, fintechs are still companies focused on being sustainable and profitable businesses that continue to attract investment, provide returns, and grow over time. As aforementioned, many of my interviewees are leaders of fintechs with backgrounds in banking and microfinance, as well as experience with entrepreneurship. This positionality is reflected in the profit motivations interviewees discussed. Chris, based in Europe and leading a fintech company that provides ML-based credit assessment models to organizations globally, noted: “Objectivity, realism, execution. That’s my mantra. Can we execute well, can we make money doing it?” Algorithmic-facilitated lending creates a seductive marriage: do good and make money.

Taken together, fintechs carry values and perspectives that are deeply rooted in Silicon Valley culture. Namely that AI technologies are objective, data is the truth, and more data is better. Furthermore, many fintechs emerge from a dual motivation: a commitment to enhancing financial inclusion and a drive for profit. This purpose-driven ethos, common in Silicon Valley and supported in the international development industry, aims to “do good” while simultaneously prioritizing financial returns. These findings reflect how algorithmic technologies are socially shaped by the values and contexts of their developers, an idea further illuminated by the lenses of technofeminism and situated knowledges.

5.4. Findings (B): A gender picture emerges

These logics inform how fintech developers perceive bias and operationalize fairness, including leading fintechs to build models that are gender “blind” or agnostic, viewing this approach as a solution for avoiding bias. The underlying belief is that making algorithms “blind” to sensitive characteristics and demographics (such as gender and caste) is the best way to ensure impartiality in credit assessments. This means the model does not consider variables of gender (or other demographics) in its credit assessment algorithm.

5.4.1. Gender "blind" & agnostic

Having algorithms be "blind" to gender was discussed across fintechs as a method to avoid bias. This was consistently brought up by organizations. Amir noted: "We try to be unbiased to the core. So, we don't collect gender, we don't collect religion, we don't collect race information." The idea is to keep the models from knowing demographics of the person and focus on the variables tied to creditworthiness. Suraj, based in California whose fintech has global operations, asserted that being unaware of gender is important for fairness and ensuring human judgment doesn't get in the way: "If you think about gender equity, I don't think we take that into consideration. Actually I know we don't take that into consideration, like male or females. That's very fair in that way… we try as much as possible to not bring gender to, you know, crowd our judgment or to make us make a decision in a different way." This belief that not considering gender means the model is unbiased to gender is commonplace. This idea is further echoed by Sarah, a fintech leader based in East Africa:

> "Gender doesn't matter in our scoring… We haven't looked at it from a gender lens because we try as much as possible not to bring in gender to crowd our judgment… Because I'm not using gender in any of our scoring, it means anyone coming onto our platform has a very level playing field with anyone else. You don't have to think about it at the back of your mind, what if?"

Gender is seen as something that could impact judgment of individuals or machines, which would not be desired as it gets in the way of the objectivity of the machines.

This idea of keeping the model objective and avoiding bias by not "seeing gender" or other demographics is tied to the logic of "leaving it to the machine". Chris captures this connection: "[We are] completely gender agnostic, color agnostic, race agnostic. This is just a statistical algorithm, machine-based learning that looked at it and then they constantly refined it… It's completely agnostic. We're just pulling data. It brings a real fairness and lack of prejudice to it." Here, Chris almost reduces the role of the fintech company and its human employees, implying that it is really the machine and its statistical algorithms that are making decisions and informing the direction. The role of humans is then to acquire the data to feed the machine and then follow the decision from the machine.

Beyond the belief that machines are objective, in many countries globally there are regulations protecting certain demographic groups from discrimination in lending. In regards to gender, 50.3% of countries have laws prohibiting discrimination in access to loans based on gender.[13] In countries with large markets for ML-based credit assessment tools including Kenya and India, it is illegal to use sex in loan decisions and therefore, lenders are not able to consider gender and other demographics as variables in ML and credit scoring processes. Given regulatory risk, the approach to be gender "blind" can also come from the top. Kumar, a data scientist based in India at a firm founded by an American, reflected, "We made a conscious effort to not bring up this [gender] factor. Especially I think the company culture… the board was pretty stern on this." Not incorporating gender was tied to concerns about risk to the company and transcended into organizational priorities.

While there is general sentiment that gender is not incorporated in the model, there are differing sentiments and strategies around tracking gender overall. Fintechs may track gender for various purposes, such as Know Your Customer (KYC) verification[14] in India, or for tracking demographics of borrowers. Others maintain that they do not know gender at all, whether incorporating it in the model or tracking it. Chris responded as follows when asked if his company tracks gender: "I don't. I'm not interested by gender. It's not my business… What I'm interested in is the number of loan applications that we score and how many scores go out. Those two things. And now I'm interested in the default rate on the portfolio. But all the other information I absolutely stay clear of. Because it can bring bias into the game." In

[13] Calculation conducted using data from Our World in Data (University of Oxford): https://ourworldindata.org/grapher/discrimination-access-to-credit-gender.

[14] KYC is a process that financial institutions and fintechs must follow in India to verify the identity of their customers before providing financial services.

addition to a method to avoid bias, this is a way to avoid risk: the fintech can claim they didn't know if there was discrimination as they don't track gender or other demographics.

5.4.2. Gender can be learned by the algorithms

While fintechs are not incorporating gender into their algorithms, there is acknowledgement that the models can pick up on correlations and learn different demographics of people, including one's gender. The idea that machine learning tools can "learn" demographics has been shown in other examples including in healthcare (Yang et al., 2024) and hiring (Dastin, 2018). Several interviewees noted that models could pick up on signals to learn gender, even if gender is not in their models or explicitly identified. In response to whether gender can be learned by the machine and it informs the machine, Gary, who is based in the US leading a fintech company with operations in Asia, responded: "It definitely gets learned. It definitely learns from that." How then does gender get learned?

There are different variables and features models can learn from to infer gender. Variables and features can be gendered, resulting in the tools learning gender from those variables and features. In regards to ability to pay, variables such as income and employment status and sector are gendered. Women face higher unemployment rates than men (4.5% globally compared to 4.3% for men) and are more often in informal work, with 80% of new jobs created for women are within the informal economy, whereas for men this number is 66% (World Economic Forum, 2023). Relatedly, up to 92% of working women in low-income countries are in informal employment, versus 87% of men (OECD & ILO, 2019). Income differences are also pronounced, including women being more likely to live in poverty as compared to men. Gender gaps in poverty are highest among those aged 25 to 34, with women in this age range being 1.2 times more likely to live in extreme poverty as compared to male counterparts (UN Women, 2023). These gender gaps are linked to persistent inequities, including – but not limited to – women overrepresented in informal sectors, institutional barriers including workplace discrimination, caretaking norms resulting in women spending more time on unpaid care than men, greater representation in part-time work, gender norms linked to traditional gender roles of men as breadwinners and women as homemakers, and more (UN Women, 2023). Another interviewee acknowledged how gender could be learned based on social networks and model features that utilize such information.

Fintechs are aware of and acknowledge how certain variables can be gendered. Kumar explained: “We don’t train them on gender, but then there’s a lot of correlated variables when you look at gender… Income is one of them. Like in India, I think [women earn] probably 85 cents for $1 that we [men] earn, right? … But more than that, the sort of employment segment matters a lot.” Another interviewee, John, noted that gender could likely be reverse engineered from the features used by the ML model. He used this to question why gender should be incorporated as a feature, because if the model knows gender regardless, there is not a need to incorporate it as its own variable.

Several interviewees acknowledged that the learning of gender may result in the algorithms leaning towards men as borrowers. This is connected to the factors discussed prior, including the greater likelihood of men being in formal employment and having higher incomes. Prashanth, based in India who is in a leadership role at a fintech with global operations, noted that gender is not incorporated as it is illegal to do so in lending (in the case of India), but, “At a technical level, you know that can still happen, which is not avoidable. It’s a preference towards a more stable segment.” Relatedly, Kumar reflected: “I’m curious about if the tools are more often granting loans to men, which makes sense based on all these different things that exist, right. And just like patriarchy, they [are] kind of learning that men are better to give loans to over time.” This perception is linked to the aspect of creditworthiness related to “ability to repay”, as opposed to “willingness to repay”. In particular, it is picking up on gendered differences related to income, job stability, and formal employment that inform “ability to repay” algorithms. Fintechs do not necessarily think that the gender differences in ML-based credit scoring and loan disbursal are a bad thing. After all, there is a sense they are accurate in predicting creditworthiness.

However, the perceived accuracy of these outcomes is complicated by the opacity of machine learning models, which also makes it difficult to determine how gendered patterns are being learned and reinforced. ML algorithms tend to operate as “black boxes”, whereby they are incredibly difficult to decipher, even to their creators (Gryz & Rojszczak, 2021). While inputs and outputs of an ML system may be known, the decision-making processes of ML algorithms (particularly those with massive amounts of data that are more complex and have higher predictive accuracy) cannot often be mapped out or understood by humans. This can be a problem because developers will not fully understand why the machine is making certain decisions or predictions. Rohan, based in India who is in a leadership role at an Indian fintech noted they were using “black box” models and explained, “We really do not know independent variables that are impacting us as much… As you go more towards black box kind of models, your ability to weed out biases becomes much more difficult… [The models can] take in patterns which may or may not be truly relevant.” Fintechs generally accepted the “black box” nature as an acceptable limitation of operating in this space.

This “black box” is acknowledged as a tension inhibiting the pursuit of fairness or equity. Kamal said, “Obviously we want to be equitable, right? But typically the models are black boxes.” In this case, the black box nature of ML algorithms are presented as a reason for not being able to tell if the models are equitable. This is a problem because models can then veil discriminatory effects over time without developers being aware. Aditi, based in the US and who works for a funder of innovations in this space, captured the concern: “Women’s economic opportunities [can actually] worsen if AI tools are introduced without care. And it can… happen in this kind of veiled way, which is the biggest sort of threat, right?”

5.4.3. Enhancing financial inclusion & gender impacts

5.4.3.1. Financial inclusion overall

Among the fintechs interviewed, the reach of customers varies, with some interviewees at smaller fintechs sharing that they reach tens of thousands of people, while those at larger fintechs highlighted reaching millions of people. One interviewee claimed to have reached tens of millions of people to date, with several million monthly active customers. Fintechs are not necessarily tracking how many of their borrowers are considered "new to credit" or unbanked. However, several interviewees offered estimated percentages of customers that are new to credit. An interviewee in India estimated that 20% of their borrower population would be considered new to credit, which equates to about two million people. These numbers can vary based on country location, as credit scoring and formal financial access varies across countries. They also can vary based on the risk appetite of firms and their investors.

Overall, interviewees affirm that AI and ML is very effective for predicting creditworthiness and repayment for unbanked and underbanked people – and continues to get better. The underlying logic is that ML tools use data that does not rely on formal access to finance to offer credit scores and loan options. Thereby, fintechs are enabling people who have been outside of financial folds to potentially access loans. This allows them to underwrite people who have not had the opportunity prior. This is affirmed in studies that have found that ML-based credit assessment models in the US can result in higher rates of credit approvals or lower interest rates for underserved consumers when compared with traditional credit scores (Di Maggio & Ratnadiwakara, 2022; Jagtiani & Lemieux, 2019).

5.4.3.2. Reported gender differences in lending

Several interviewees did not know, or were not willing to share, the gender breakdown amongst users. In response to questions about gender differences in who gets access to loans through the app, Kumar said: "Generally we haven't looked into this… I'm not sure whether anyone on the team has. This has never [been] brought up." Kumar was skeptical that I would get data about borrower demographics from any fintech company. While no interviewees provided raw data of the gender breakdown of borrowers and loan sizes (despite my asking for this data at the end of the interviews), over half of the fintechs interviewed shared approximate numbers in the interview itself. Others would be more vague noting that more men accessed loans without providing specific numbers. There was also a lack of knowing impacts of their loans, including as it relates to gender. John noted that they simply don't have time to track impacts, much less disaggregated by gender.

Across the information interviewees provided, a gendered pattern in lending emerges. First, interviewees shared that they provide more loans to men, except in a few countries where the breakdown is more balanced, and that loan sizes tend to be smaller for women. They did not clarify if this disparity is due to being offered smaller loans compared to men, taking smaller loans compared to men even if they were offered higher loans, or a combination. These gender differences are not isolated: interviewees reported similar patterns across multiple country contexts, suggesting a broader trend that extends beyond individual fintechs. However, the extent of gender gaps, both in borrower representation and loan size, varies by country – with similar country variations shared across fintechs. Josh, who works at a fintech company with operations in Kenya, Mexico and the Philippines, said, "The three different countries are all very different, so they have different ratios of… males and females." Algorithmic-facilitated lending tends to have the highest gender gaps in India followed by Kenya / East Africa, compared to other primary market locations (e.g., Mexico, Philippines). In India, Amir acknowledged 20 to 25% of borrowers are female and 75 to 80% are men, while Kumar shared similar numbers at his fintech with 30% of borrowers being female and 70% male. Other interviewees from fintechs in India would not provide specific numbers (whether because they did not know them or because they would not share them). Kamal noted the number of borrowers is "predominantly men", while Rohan said simply, "generally the male goes and applies for a loan. The women typically don't apply."[15] Sarah acknowledged it as slightly better in Kenya with her fintech having approximately 35% of female borrowers and 65% male borrowers. Interviewees with fintech operations in Mexico and the Philippines shared numbers that were closer to 50% for men and women.

[15] While finding exact numbers on the gender breakdown of borrowers and getting data from the companies was not possible, several interviewees provided numbers. I had listed the question in my interview questions which were sent prior (so they could prepare as desired) and always asked if further data could be shared, but despite some initially agreeing, did not receive any of the written data.

Although a limited sample, these trends track with gender inequality rankings – meaning that the countries with the larger gender gaps in ML-based lending correlate to countries with higher gender inequality. The Global Gender Gap Index assesses national gender parity and ranks 146 countries. In this, India is listed near the bottom at 129, Kenya in the middle at 75, Mexico near the top at 33, and Philippines nearby at 25 (World Economic Forum, 2024). These trends also track and are linked to gender digital divides, with India having the highest gender digital divide, followed by Kenya, and Mexico and Philippines being closer to parity (Digital Gender Gaps, University of Oxford, 2024[16]).

Importantly, trends related to gender gaps in algorithmic-facilitated lending are not set in stone. Some fintechs target specific groups that hold gender implications. For example, an interviewee at a fintech company that lends to farmers, discussed how the gender of farmers differs in their two country locations. In one market location, Pakistan, almost all farmers who are in charge of and running the farm are male. In the other market location, Thailand, farmers who are running the farms are predominantly women as more men work in factories. So they predominantly lend to women in that country. One interviewee emphasized the importance of being gender intentional. Rohan, who is based in India and has worked at two different fintechs in the ML-based credit assessment space in India, noted: "On the gender side, it has to be a conscious effort. You will gravitate towards giving loans towards more male borrowers. So unless you want to make a conscious effort to say more women, it's unlikely naturally for it to become much more gender equivalent." This illustrates that decisions (including who to target and where) have gender implications.

5.4.3.3. Women considered better repayers

Across the board, interviewees reported that women are better repayers, including that they are both more likely to repay on time and to not default. Interviewees acknowledged this as a pervasive trend and offered reasons linked to observations about the trustworthiness and reliability of female borrowers,

[16] https://www.digitalgendergaps.org/

indicating that this trend remains regardless of loan size. Interviewees acknowledge that from a business perspective, this makes them a better bet. Kamal said: “Women are proven to be better repayers so from a business point of view it makes sense to actually have women. But that’s not what people do.” Sean, who is based in California and is a data scientist leader at a fintech with credit assessment operations largely in sub-Saharan Africa but with intentions to expand more broadly, echoed: “Women are getting the smaller loans on average, but they are also far better at paying them back. Women are just a much better bet if you are a bank… You want [to lend money] to the women, not the men.” Nicholas, who is based in Europe and is a leader at a fintech with global operations, noted that this is a global trend they have witnessed with women being better repayers in every country location they have worked in. This was further echoed by Dante: “There are some local rules that apply globally. If you are a woman, you are a better repayer than a man. Full stop.” This tracks with evidence that suggests women tend to be better at repaying loans, particularly in the context of microfinance.

Growing evidence suggests women tend to be better at repaying loans in the context of microfinance. Research assessing repayment performance of microloans in Bangladesh, Guatemala, Malawi, Malaysia, and Mexico show that women have better performance (Armendariz de Aghion & Morduch, 2010). A study using a global dataset of 350 microfinance institutions in 70 countries confirmed that a higher percentage of female clients in MFIs is associated with “lower portfolio risk, fewer write-offs, and fewer provisions, all else being equal” (D’Espallier et al., 2011). MFIs and multilateral agencies have thereby targeted women, asserting women as good credit risks. Case studies from other lenders in countries with sex-disaggregated supply side data, show that women have lower default and higher savings rates than men (Buvinic & Ruf, 2022). While research illustrates women as better repayers relative to men, D’Espallier et al (2011) also note that focusing on female clients is associated with smaller loans and lower Operational Self Sufficiency. In other words, women may not generate as many financial returns as male clients, putting into question the business case for focusing on female clients. There is limited research regarding *why* women may be better repayers than men. One study conducted a trust game and a microloan repayment game in rural Bangladesh to find that women are more trustworthy than men and more likely to repay loans irrespective of other conditions, such as joint liability or repayment incentives (Shahriar et al., 2020).

Regarding why women are considered better repayers in algorithmic-facilitated lending, there are various reasons discussed by fintechs. These include women being regarded as more financially disciplined or financially responsible, including related to the reasons for taking a loan. Rohan explained: "Women borrowers are more financially disciplined and want to leverage it for the right purposes… That's something that we see over and over again… Once you give [a loan] to women borrowers, it is most of the time used for the right purposes." This links to the aspect of creditworthiness regarding intent (or willingness) to repay. Various interviewees discussed that men are more likely to spend money on things such as alcohol, versus items for their family, and that men can be less disciplined financially as compared to women. Amir, based in India, discussed that between 20 and 25% of their borrowers are women but "in terms of repayment, they are far superior… I think they have better financial control… financial discipline is very high in female borrowers." Others discussed that women are more responsive to collection calls than men. There is also an aspect of psychology. Rakesh said, "What we've noticed generally is that women are way more responsible, conscientious, willing to change, adapt, learn than their male counterparts overall." This trend highlights a latent tension exists around stated ideals of opening new roads to finance in meritocratic ways and acknowledgement that women are better repayers. Persistent barriers prevent stated ideals from being met. It also brings up questions around why there are enduring gender differences in algorithmically-facilitated loans.

5.4.4. Reasons for gender differences in loans

In this section I outline the observations by and perceptions of fintechs regarding the reasons for gender differences in algorithmic-facilitated lending. I apply socio-structural explanations (Smart & Kasirzadeh, 2024) to these observations and perceptions in order to illustrate the ways different factors are gendered. I discuss how algorithms may be optimized by fintechs and why this matters, before outlining gender implications and a key tension that remains.

5.4.4.1. Perceptions regarding why there are gender differences

Interviewees outline several observations regarding why there are gender differences in ML-based credit assessment and loans. At a high level, interviewees highlight that gender differences related to loans can be linked to broader gender norms, structural inequities, and cultural differences. This manifests in several demand-side factors. First, interviewees note that there tend to be less women applying as compared to men. Some acknowledge that this may be due to less access to apps given the medium of the tool being smartphones and gender differences in mobile ownership. Other interviewees noted that women are more likely than men to have basic phones, or use a shared phone. Sarah explained the situation in Kenya: "Being a patriarchal society, if there are two handsets in a home, one more superior than the other, the superior one will belong to the leader of the household." Interviewees acknowledge that this factor is more relevant for rural areas where gender differences in smartphone penetration and Internet tend to be more pronounced. Relatedly, interviewees reflected that less women may choose to apply for loans. Some noted that the app may attract more male users, given their role as household financial managers. This is captured by Rohan: "If a person needs a loan it is generally the male who kind of goes and applies for a loan… Typically you would see the skew towards more men." Third, even if they do apply, interviewees noted that women tend to face greater challenges using the app compared to men and require more support. Taken together, these factors highlight ways that fintechs have observed demand-side constraints for women in accessing loans.

From the supply side, interviewees reflected on the ways creditworthiness is defined and operationalized in the machine learning tool. Several discussed how variables related to employment can have gender implications. Rohan noted: "It's more again a social thing… In general men are more employed compared to women. There are more women who are, you know, homemakers and other things in India, compared to men. So that's one piece which kind of skews." Kumar discussed how predominant roles of women as household caretakers impact ML-based credit assessments. He said: "We don't discriminate against [women], but then definitely employability segment is one… So maybe housewives would probably be in a less fair situation to get [a] higher ticket size… These sorts of things will end up coming in." This reflection illustrates gendered understandings around features used by algorithms to predict creditworthiness, particularly related to *ability to pay*.

5.4.4.2. Applying socio-structural explanations to perceptions of gender differences in loans

Applying socio-structural explanations to make sense of algorithmic decisions is an approach that can improve the scope of interpretations for model operations. ML models do not operate in isolation, but within complex social and institutional constructs that can significantly impact their behavior and impact (Smart & Kasirzadeh, 2024). This approach allows me to expose how ML-based credit assessment apps are linked to and embed gender norms and structural inequalities.

The demand- and supply-side factors observed by fintechs are linked to gender norms and structural inequalities. First, fintechs observe that less women are applying to the apps and note that women may have less access to the apps given the medium of the tool being smartphones. Due to gender digital divides less women have access to smartphones and Internet required to access apps. There are persistent global gender differences in digital inclusion and literacy (i.e. the gender digital divide) as well as financial literacy (Hasler & Lusardi, 2017). Surveys and research by GSMA reveal that across LMICs, women are 8% less likely than men to own a mobile phone and 20% less likely to use the Internet on a mobile, with gender gaps further amplified in rural versus urban areas (GSMA, 2024). In two of the most popular countries for these apps (Kenya and India) gender gaps loom: in Kenya there is a 34% gender gap in mobile Internet use, whereas in India the gender gap is 52% (GSMA, 2024). Unsurprisingly, there is a fintech gender gap: in a global study of 28 countries, 21% of women use fintech products compared to 29% of men (S. Chen et al., 2023).

Gender differences in digital inclusion and literacy, as well as financial literacy connect to the other observation by fintechs in which women can face greater challenges using the app compared to men. Linked to persistent gender discrimination and limiting gender norms, women – and those with lower socio-economic status – are more likely to need more support to use financial accounts (World Bank, 2021). Important to acknowledge is that women are not a monolith. While many women have access to the Internet and smartphones, as well as strong digital and financial literacy, the proportion of women with this access and literacy is lower as compared to men, particularly in rural areas.

Secondly, interviewees reflected that less women may choose to apply for loans. This aspect of choice can be linked to gendered household financial decision-making. In many countries, bank accounts are more often in the name of the man of the household. This is reflected by gender differences in account ownership globally. Women's account ownership is 6 percentage points lower than men's in LMICs (World Bank, 2021), with gender gaps that are further inflated in rural and lower income groups. If loans are distributed to bank accounts, as some fintechs operate, the loan may not be provided to the women. Also, women may not choose to apply or choose lower loan sizes linked to lower tolerance to financial risks. While not necessarily "good" or "bad" in the context of ML-based lending apps, research illustrates lower levels of tolerance to financial risk given managing family risks and may avoid high risk financial options (Croson & Gneezy, 2009).

Looking at the supply side, the ways creditworthiness is defined and operationalized in the ML-based credit assessment tool can favor men inadvertently, linked to gender norms and inequalities in a country. Interviewees acknowledge the gendered nature of certain features or proxies that algorithms consider, particularly regarding algorithms assessing *ability to repay.* These features or proxies (e.g., income level, consistent or stable income, formal employment, and cash flow) hold and reflect gendered differences and structural inequalities in economic status, employment, and caretaking responsibilities. Women tend to have lower incomes as compared to men, be in more informal employment (versus men as more often in formal employment), have less consistent or stable incomes as compared to men, and tend to have higher caretaking responsibilities or roles as homemakers (versus men as more household financial managers and decision makers) (World Economic Forum, 2024). In assessing cash flow, which is a common proxy within algorithms assessing ability to pay, an assessment made on an individual woman's phone may not be reflective of the broader cash flow of her or her family. Rather, if household finances are managed by men in the household, key information could be missing from transactions documented on a woman's phone. Differences in how women and men use their phones, which are linked to gender norms, may also influence the algorithms in other ways that are not fully understood.

Building from the observations and perceptions of fintechs, a picture forms regarding the ways ML-based credit assessment apps are linked to and embed gender norms and structural inequalities. It is important to acknowledge that gender gaps in algorithmic-facilitated lending vary less between fintechs and more between country locations, as aforementioned: interviewees report greater gender gaps in algorithmic-facilitated lending in countries with greater gender inequality and vice versa.

5.4.4.3. What is optimized for matters

There's a key consideration between predicting creditworthiness and optimizing for profitability. Some interviewees shared that ML-based credit assessment algorithms don't just assess risk — they also estimate the "lifetime value" (LTV) of a customer. LTV, a common metric used in lending, reflects expected revenue from interest, fees, and long-term engagement, minus acquisition and servicing costs. Two large fintechs noted that LTV is a key component in their credit assessment. Megan, who is based in California and works at a fintech company with global operations, noted: "You can get to predict lifetime value… And that's how the business can make decisions, like what is our payback window? How much are we worth? Are we willing to wait to get to a net positive return on this user? That's how we make decisions on approval." By incorporating the potential level of returns over time, assessing creditworthiness begins to reflect more than just one's ability to repay and willingness to repay. Now, factors like how much one may borrow over time, at what interest rate, and with what level of fees or penalties become part of the decision. In this, business priorities, particularly profit optimization, shape decisions. These choices are often not framed as value judgments, but rather good practices for financially sustainable lending. Yet, this framing can obscure how deeply commercial logic is embedded in decisions that affect access to credit even under "for good" umbrellas.

Consideration of profitable loans and borrowers is not new to lending. Indeed, lenders, being in the business of making money, tend to prefer asset wealth clients who can take larger, and thereby more profitable, loans (Buvinic & Gokhroo, 2023). Gender differences can be present: among fintechs that do not tailor products to women, only 38% report that women's LTV is higher compared to men's (however, this jumps to 63% for the few fintechs that customize products and services for women) (International Finance Corporation, 2024).

In the case of ML-based credit assessment, optimizing algorithms for lifetime value, may be a contributing factor leaning towards men in credit scoring (who tend to have higher incomes, take higher loans, and reportedly have more late payments that can generate greater returns). This would not be the first documented case of ML tools in lending exhibiting a bias due to optimizing for profit: Research on a peer-to-peer lending platform in China found that the introduction of ML to inform interest rates resulted in higher interest rates for women. This was not due to a higher estimated risk or lower determined creditworthiness by the algorithm, but because women had lower price sensitivity and the platform could therefore better optimize revenue by offering women loans at higher interest rates (Chu et al., 2023). Algorithms optimized for profit may inadvertently penalize women – not because they are riskier, but because they may be less profitable under lifetime value models. In this research, interviews report higher loan sizes and fee-generating behaviors, such as late payments are more often associated with male borrowers. These behaviors have the potential to contribute to higher lifetime value scores and may help explain gendered differences in lending outcomes. I discuss and explore this further in Chapter 6.

5.4.4.4. Gender implications & perceived responsibility of fintechs

While fintechs recognize the role that gender norms and inequities play, they do not thoroughly understand and/or account for gender differences in access and use of phones, as well as ensuing implications for algorithms. Hannah asserted: "I mean a lot of it is just the technology intersecting with the existing norms, culture, and society that exists. Then maybe for the people building those tools there isn't as much understanding of how it might be different for women." No interviewees expressed an intention to perpetuate gender inequities. Rather it is seen as an unintentional outcome, but is justified because access to finance provided can still be better than the status quo.

5.5. Discussion

5.5.1 Algorithms reflect their creators

My findings reveal various logics that inform algorithmic design and managerial decisions, revealing how knowledge is situated. Algorithms are results of the perspectives and priorities of creators, which are shaped by cultural and institutional contexts and logics. This aligns with a central technofeminist insight: technological design is not neutral, but deeply influenced by the social positions and contexts of those who build it. My findings reveal two dominant mindsets shaping algorithmic development, one rooted in American banking and the other in Silicon Valley.

American banking logics influence the conceptualization, design, and management of ML-based credit assessment tools including through conceptualizing creditworthiness in terms of personal wealth and one's individual character and trustworthiness. This approach to credit assessment and provision is not universal. In other alternative credit systems, such as savings and credit cooperatives or microfinance institutions, creditworthiness is not narrowly defined by individual financial behavior or moral judgments, but rather collective responsibility, social ties, and consistent participation (Lee & Persson, 2024; Postelnicu & Hermes, 2018). These alternative approaches illustrate that credit assessment and provision can be grounded in community dynamics as opposed to individualized, meritocratic assumptions and data-driven notions of character.

ML-based credit scoring tools reinforce and repackage the ethos of American banking by defining creditworthiness in terms of ability and willingness to repay. Assessment of one's character, morals, and trustworthiness is central: will they *choose* to repay, not just *whether* they can repay. There is an inherent value of meritocracy, in which anyone can lift themselves up with hard work, commitment, and strong character. There is also a subjectivity attached to assessing what is sufficient for one's ability to repay, including preferences for formal versus informal jobs, as well as in assessing one's character. Within ML-based credit assessment tools, American banking logics intersect with Silicon Valley logics, in which more data is better, data is the truth, and leaving decision-making to ML algorithms is best given their objectivity.

The interconnectedness between developers' cultural and institutional contexts and the design of ML tools underscores that knowledge is situated. American banking logics lead to credit scoring algorithms that include assessing one's ability to repay and willingness to repay, which is about someone's morality and trustworthiness. Meanwhile, Silicon Valley logics of believing in data as the truth and "leaving it to the machines" as objective technologies, inform choices of fintech developers. In particular, this perspective often leads fintech developers to build models that are gender "blind" or agnostic, viewing this approach as a solution for avoiding bias. This belief in the technology's objectivity ignores the ways that technology in *use* is different from technology as an *object*; whereby its use derives meaning and effects from contextual factors (O'Connor & Liu, 2024). This belief in the objectivity of technology and data – coupled with a mission to promote financial inclusion – can result in a limited critical reflection on the subjectivity inherent in the concept of creditworthiness. These logics are reflected amongst all interviewees: as in, it is not just Americans who hold American banking logics. Indeed, elites in LMICs – many of whom live in cities and are trained in Western education systems – can also be guilty of imposing Western categories and technical knowledge (Parpart et al, 2000).

5.5.2. Belief in the machine's objectivity reinforces and obscures existing gender inequities

Benjamin's concept of default discrimination highlights the gap between intention and outcome in technology design. By adopting a gender "blind" approach, fintechs can obscure existing disparities and fail to directly address the underlying structural inequalities that algorithms then learn from. As a result, ML-based credit assessments may inadvertently perpetuate gender inequities and mask the reproduction of existing power hierarchies.

Gender "blind" or agnostic models result in an inherent flaw: the lack of grappling with the ways data is gendered and reflects inequities, as well as the ways features and proxies connected to ability and willingness to repay are gendered and can be subjective. While there is some recognition that features or proxies used in creditworthiness assessment (e.g., income, stable employment) can be gendered, there is less consensus as to whether this is an issue. The tool is designed to assess creditworthiness, and if these factors are important to creditworthiness, then it is doing its job – even if it is more likely to deem men as more creditworthy. However, in general, there is a lack of consideration on how data can be gendered and lead to misleading or inaccurate predictions regarding creditworthiness.

While it may be intuitively appealing to believe that an algorithm can remain "blind" to gender (or other demographic attributes), a core issue arises if and when developers fail to recognize that proxy variables can inadvertently encode and reproduce gendered patterns – leading to discriminatory outcomes. This has been seen in algorithms in other domains. For example, a widely used healthcare algorithm in the US did not incorporate race into the algorithm nor consider how different racial groups may experience the algorithm. Researchers found that the algorithm falsely concluded Black patients as healthier than equally sick White patients, which reduced the number of Black patients identified for extra care by more than half. This bias occurred because the algorithm used health costs as a proxy for health needs and Black people spend less than white people in healthcare settings in the US (Obermeyer et al., 2019). Algorithms can also learn that certain groups are "better" and prioritize those individuals. For example, Amazon developed a hiring algorithm to review job applicants' resumes and identify top talent. The algorithm was trained on data from the company whereby men were in top technical and management positions. Although the resumes were "blind" to gender and race, the model learned that top talent was male and thereby systematically downgraded women's resumes (Miasato & Reis, 2020). The technology identified which resumes belonged to men and women, including by favoring candidates who described themselves using verbs more commonly found on male engineers' resumes, such as "executed". Unable to fix the bias, the company scrapped the tool (Dastin, 2018).

Relatedly, there is a general lack of grappling with gendered differences in regards to access to the apps, which can limit how many women apply to loans via the apps to begin with. This lack of grappling with gender differences in access to apps is not true for every fintech. One fintech company operating in East Africa, for example, recently started making ML-based credit assessment available through basic phones (in addition to their traditional smartphone app) as part of an effort to increase access to more financially excluded people, particularly women. However, by and large, there is a lack of grappling with apps being an exclusive medium for many people, disproportionately women, and not seeing that as an issue the fintech is responsible to solve. This could result in greater data on men versus women that the ML models are trained on. This matters in machine learning: if the dataset the model learns from over-represents a certain identity it will perform better for that particular identity. This has been shown in various applications of AI, such as facial recognition algorithms (Buolamwini & Gebru, 2018), medical imaging (Larrazabal et al., 2020), speech recognition (Koenecke et al., 2020), and education outcome prediction (Gándara et al., 2024).

Data is not necessarily the truth and ML models are not objective. What data an ML model is learning from matters, as it can over or under-represent different groups with implications for who it performs better (or worse) for, and data collected on a person's smartphone can differ in terms of its quality and depth. Data also carries with it a reflection of inequalities that exist in society. Meanwhile, the variables linked to definitions of creditworthiness matter.

Not addressing these gendered aspects of the technology head on is linked to three areas: a deep belief in the objectivity of data and machine learning leading to lack of awareness or recognition this is an issue; avoidance due to regulations around not being able to consider gender and other demographics in scoring and keeping distance by insisting that gender is not seen or taken into account; and a sense that it isn't the responsibility of the fintech to address structural gender inequalities. These algorithms are not neutral tools but rather products of specific cultural and institutional frameworks. ML-based credit assessment tools follow Benjamin's default discrimination, in which design processes stay "blind" to gender and ignore social cleavages and inequities, thereby defaulting towards continuing discrimination that exists in society.

By applying labels of objectivity to credit scoring via machine learning, power hierarchies and social norms are ignored and – often inadvertently – embedded under veils of objectivity. This highlights a phenomenon I call the *objective algorithm paradox*, in which the belief in machine learning's objectivity results in being agnostic or "blind" to certain demographics in ways that hide the gendered nature of algorithms and perpetuate discrimination. The belief by algorithm developers and fintech leaders that machine learning is objective obscures the moral choices they are making in design and management choices, resulting in a lack of grappling with those moral choices.

This doesn't mean that gender should necessarily be a factor in algorithmic decision-making. Rather, auditing across demographics is critical to ensure certain groups aren't being penalized, as well as conducting assessments of how datasets training models and features can be gendered to mitigate issues proactively. In the case of the healthcare algorithm, the researchers worked with the company that developed the algorithm to identify and remove the features that were serving as proxies to race (e.g., health spending), which mitigated the racial bias (Obermeyer et al., 2019). Similar efforts could be taken in the fintech lending space. More, it is critical to consider how the algorithm – even when working as intended – may reinforce inequalities that exist in society. As the tools are "black box", this reinforcement becomes particularly difficult to spot, with the potential to therefore solidify and project inequalities into the future. The goal of these technologies shouldn't necessarily be parity of equal loans to men and women (this could result in unintended consequences of giving loans to people who can't repay them). However, it is important to be transparent about the tradeoffs in algorithmic-facilitated lending and then have discussions, including with impacted consumers and marginalized groups, about what is acceptable or not, as well as how algorithms should be operationalized and considerations around fairness. The concept of fairness is further discussed in Chapter 6.

In short, technology and society are inextricably linked (Subramaniam et al, 2017). This matters in AI development. Power, privilege, and oppression are not equally distributed (hooks, 1984; Crenshaw et al., 1996); yet AI technologies, largely controlled by a "coding elite" (Burrell & Fourcade, 2021), often ignore power hierarchies and inequalities, resulting in default discrimination under veils of objectivity (Benjamin, 2019). Meanwhile, data reflects patterns of societal discrimination and inequality; thereby ML tools learning from this data fall prey to reinforcing those patterns (Eubanks, 2018b).

5.5.3 ML-based credit assessment results in gender differences in lending

While fintechs and their ML-based credit assessment tools are increasing access to finance overall, this access to finance is not gender equitable. Interviewees report providing more loans to men and loans at higher amounts. These disparities are more pronounced in countries with higher levels of gender inequality, where gender norms and structural inequities influence both the demand and supply sides of credit. Intersectionality can also play a role in differences.

Interviewees note that gender differences related to loans come from demand-side and supply-side factors, which are linked to gender norms and structural inequality. Fintechs acknowledge that women tend to apply less than men, while also having relatively more challenges in using the app. These differences can be linked to gender digital divides, gender norms around household financial decision-making, and differences in digital and financial literacy. On the supply side, interviewees highlight that features and proxies particularly used in algorithms assessing ability to pay (e.g., stable employment, income levels), can be gendered. Gender differences in algorithmic-facilitated lending reported by fintechs correlate to levels of gender inequality, highlighting how ML-based credit assessment apps reflect gender inequalities in the context in which they are deployed. Furthermore, some companies are optimizing for lifetime value of customers, which may be a contributing factor leaning towards men in credit scoring (who tend to have higher incomes, take higher loans, and potentially have more late payments that can generate greater returns).

In all of this, there is a key tension: interviewees observe that women are more likely to repay on time and not default on loans as compared to men, despite being less likely to access loans and have lower loan amounts than men. This tension represents potential discriminatory effects and a market inefficiency resulting from the algorithm, particularly linked to ability to pay algorithms. In countries such as Kenya and India where these tools are popular, due to gender norms and structural inequities, women tend to have lower income levels and are more often in informal jobs compared to men, leading to less consistent income or stable cash flows. While this may be a relevant indicator for creditworthiness for larger loans (e.g., mortgages) these aren't necessarily indicative of ability to pay back small and micro loans. In addition to representing a potential discriminatory effect, this represents a market inefficiency.

The specific mechanisms of gender differences are largely hidden by algorithms that are closed to external evaluation, as well as the "black box" nature of ML itself. Even if algorithmic-facilitated lending is providing more loans to people than the traditional status quo financial system, ML presents the potential of solidifying gender inequities in ways that are not fully understood and projecting them into the future under veils of objectivity. They could also be amplifying gender inequities given the non-linear nature of ML models. This exploration illustrates how default discrimination can manifest in these contexts: fintech companies may acknowledge the existence of structural inequalities yet still adopt a gender "blind" approach in their algorithmic design, follow other priorities, and ultimately, not tackle inequalities in how they design and manage algorithms.

5.6.4. Encoded gender norms

I build from prior theories that examine algorithmic inequality, bias, and power broadly and/or in relation to race, to introduce the pattern of *encoded gender norms,* in which the status quo is solidified. In this, there are several key propositions: (1) Data and features are gendered (as is access to technological tools); (2) not considering the ways data and features embedded in algorithms are gendered can replicate and reinforce gender norms (this gender ignorance is not bliss, but rather problematic for gender equality and firms alike); and (3) prioritizing inclusion while also being gender "blind" comes at a cost of equity. This leads to algorithms encoding gender norms in ways that result in self-fulfilling prophecies that become harder to spot and solve. While gender norms can and do evolve in and across societies over time, ML tools may impede this evolution under opaqueness and veils of objectivity. As ML tools serve as mirrors, we can expect greater gender variance depending on gender inequality levels, while greater attention will need to be placed in gender unequal areas before potentially risking reinforcing and legitimizing limiting gender norms and inequalities. There is opportunity to disrupt the encoding of gender norms through gender intentional and transformative processes and design.

5.6. Conclusion

This chapter outlines the underlying logics and mindsets informing the development of algorithmic lending tools (findings A) and how those lead to choices around gender "blind" algorithms and implications for perpetuation of gender inequities in financial access, despite women being observed as better repayers (findings B). The idea that technology is inherently objective and, when left alone, will be fair, is both seductive and misleading. In reality, knowledge is situated and algorithms are instruments of values. They reflect the values of the people and – more – the institutions that create them. Believing in the objectivity of machine learning creates and obscures unintended gender bias in the ML-based credit assessment with resulting gender differences in lending and perpetuation of gender inequity as the default. While fintech employees and investors are focused on "doing good", in the context of algorithmic-facilitated lending in LMICs, decisions and tradeoffs ensue. Ironically, that women are better repayers but not getting similar loans through current approaches to algorithmic-facilitated lending illustrates that the promise of meritocracy pursued under American banking logics is not achieved.

This prompts the questions: What do fintechs perceive is fair? Where do these perceptions come from and how do they influence the ways gender differences in algorithmic facilitated decisions are legitimized (or not)? The next chapter focuses on this by exploring differing definitions and approaches to fairness. It delves into the ways capitalistic priorities of for-profit financial institutions embed particular perceptions and measurements of fairness that influence credit assessment and lending.

CHAPTER 6: PERSPECTIVES AND PRIORITIES – WHAT IS FAIR?

6.1. Introduction

John, the American CEO of a fintech employing ML for lending in LMICs, knows that women receive fewer loans and loans at lower amounts in the markets where his company operates. He is dedicated towards his company's social impact mission of enhancing financial inclusion, while also identifying as a businessman in a tough industry. He remains frustrated by the "social justice warriors" whom he feels demand that loans should be distributed equally and that lenders must address economic injustice. When asking about fairness, particularly after discussing gender differences in lending outcomes, I seem to have struck a nerve. Though he finds it interesting to consider fairness having studied philosophy in his graduate education, it stirred irritation and he reflected that he's found himself turning away from the topic. His experiences and emotions expose the complexity of ethical decisions that underlie choices in AI model design and management.

This chapter extends the analysis developed in Chapter 5 by shifting from the implicit assumptions guiding technology design and management (i.e. being gender "blind"), to the explicit ethical frameworks fintechs use to justify, inform, and evaluate their technologies, particularly the concept of fairness. While Chapter 5 showed that developers rarely account for how gender is embedded in data and design and outlines fintech reasoning for gender differences in lending, Chapter 6 deepens this critique by analyzing how fintech actors define what counts as "fair." This matters because fairness functions as both a technical goal and moral justification in AI development and management. By unpacking its varied interpretations, the chapter reveals how dominant definitions can reinforce gender inequalities even as they are framed as ethically sound.

Drawing on grounded theory and inductive thematic analysis of the same 25 interviews, this chapter examines how fintechs define fairness, including how perceptions of fairness reinforce gender "blind" approaches and legitimize gender differences in algorithmic-facilitated lending. I explore how fairness perceptions are shaped by underlying American banking and Silicon Valley logics, as well as pursuit of "accuracy". This matters because these definitions do not just shape technical models, but (as I will demonstrate) also influence how their outcomes are evaluated – informing or justifying both the processes fintechs adopt and the benchmarks by which they judge success. By embedding narrow or decontextualized views of fairness into their systems, fintechs risk reproducing structural gender inequalities while maintaining an appearance of objectivity or neutrality.

The contribution of this chapter is three-fold. First, it sheds light on how fintechs' definitions and considerations of fairness are value-laden and shape decisions in the development and management of algorithmic lending tools, while failing to challenge gender inequities in access to finance. Second, it builds on and pushes beyond theoretical debates about the role and limitations of AI fairness in algorithmic accountability by using empirical evidence to critically demonstrate and examine how fairness is operationalized in the real-world context of fintechs that are leveraging "AI for Good". Third, it contributes directly to my RQ1 by showing that, across diverse fairness framings, ML-based credit assessment tools consistently fail to challenge gender inequities in access to finance – revealing how structural inequality is reproduced, not mitigated, through current fintech practices. In doing so, it reveals the normative decisions made by developers and managers of the technologies that inevitably intersect with tensions, contradictions, and power dynamics embedded in these efforts.

This chapter is organized in several sections. I begin with an overview, drawing on the larger literature, of why fairness is a critical topic to examine, different disciplinary understandings, and how fairness is operationalized in ML technologies in order to ground the paper and its findings. After explaining my methods, I present how fintech leaders perceive fairness in the context of algorithmic lending. This examination considers fairness from both a process perspective (e.g., what constitutes a fair approach in designing ML tools) and an outcome perspective (e.g., what is deemed a fair result from the application of ML in credit assessments). I conclude with a discussion on my findings.

6.2. Background on fairness in machine learning

While the core literature review focused on gender bias in AI, as well as impacts of ML-based credit assessment, this chapter's background section delves into the algorithmic fairness literature. This literature extends and complements the broader bias in AI literature. It provides more granular understandings of how fairness is defined, measured, and operationalized in algorithmic systems. The relevance of this algorithmic fairness literature emerged through the findings for RQ1, in which fairness surfaced as a key aspect in the underlying logics of ML-based credit assessment design and management. This reflects the iterative nature of the qualitative research process.

6.2.1. Why is it important to understand fairness and what does it mean?

AI and ML tools are increasingly governing more aspects of our lives – in finance as well as healthcare, education, and more. These systems make predictions (e.g., if one is creditworthy) while optimizing for specific objectives (e.g., maximizing a certain definition of accuracy and/or profit). Fairness is not inherent in ML, but must be deliberately defined and encoded into the system. When fairness is not explicitly prioritized, ML systems are left to optimize freely for the goal they are given, identifying and reinforcing patterns in data without questioning whether those patterns reflect structural inequalities. This can lead to outcomes that systematically disadvantage certain groups, such as women, while still appearing neutral or technically sound.

Across domains, there are pervasive fairness concerns and recognition that it is important to ensure models are fair (Barocas & Selbst, 2016; Pessach & Shmueli, 2022). But what does fairness mean? Fairness is commonly defined as the "quality or state of being fair, especially fair or impartial treatment"

(Merriam-Webster, 2025). It is a ubiquitous, yet contested, term in AI and ML research and practice. After all, the state of being "fair" or impartial, remains unclear and nuanced.

In ML applications like lending, fairness takes on especially high stakes. Developers must make normative choices: should a lending model aim to maximize profit, increase the number of approved loans, or ensure equitable or equal access across groups? These choices are not merely technical and they directly invoke questions of fairness with ensuing impacts on how the algorithm learns and operates. The choice of fairness definitions and approaches has immense social impact consequences in lending (Liu et al., 2018). Misclassification costs individuals dearly by denying loans that could have been repaid resulting in missed opportunity for the individual to enhance their social or economic position; or receiving a loan one cannot repay may worsen one's financial situation (Kozodoi et al., 2022).

Fairness is not a universally agreed-upon concept and has different meanings in various domains (Mulligan et al., 2019). Legally, fairness relates to anti-discrimination and protecting individuals or groups based on protected characteristics. In the social sciences, it is understood through the lens of power dynamics, institutional arrangements, and structural inequality. In quantitative fields (e.g., computer science), fairness is framed as a mathematical optimization problem, using formal criteria such as equal error rates, representation, or allocation to assess outputs. Philosophical perspectives often tie fairness to broader principles of justice and equity.

Fairness can also mean different things in different contexts to different people. In his book, *Voices in the Code,* David Robinson explores a Kidney Allocation System, which uses an algorithm to identify who should be prioritized for new kidney transplants in the United States. The system incorporated perspectives from a variety of stakeholders – including patients, surgeons, data scientists, public officials, and more – to collaborate on the development of the algorithm including by identifying what should be considered fair. Robinson highlights debates that occurred from various stakeholders about what the algorithm should be optimized for and what was most fair, including questions related to utility versus equity. Perceptions of fairness differed for different stakeholders. Moral choices and ensuing decisions around fairness were both necessary and consequential for the design and operation of the algorithm (D. Robinson, 2022).

6.2.2. How is fairness in machine learning measured and operationalized?

In ML practice, fairness is often approached as a quantitative optimization problem, focusing on constructing an "optimal ML model subject to fairness constraints" (Mulligan et al., 2019). These constraints introduce formal rules into the model training process, requiring the algorithm not only to perform well overall but also to satisfy specific fairness criteria. In domains like credit scoring, models predict whether an applicant is creditworthy and likelihood of default (e.g., through a score). These scores are typically converted into binary decisions (e.g., whether to approve or deny a loan). Model performance is therefore commonly evaluated using a confusion matrix, which breaks down predictions into true positives, false positives, true negatives, and false negatives (Mirpourian et al., 2022). This structure allows developers to calculate fairness metrics (e.g., false negative rates) across demographic groups. These metrics form the basis for applying fairness constraints. For example, a model may be trained to minimize disparities in false negative rates or ensure equal true positive rates between groups (Mirpourian et al., 2022).

Before fairness constraints can be implemented, developers must decide what fairness means in a given context and how it should be measured. Conceptual frameworks provide high-level guidance on what fairness means in practice, often by specifying how to balance outcomes across groups. In finance and credit scoring, three common conceptual frameworks are often discussed (Kozodoi et al., 2022): *independence* ensures equal loan acceptance rates across groups, but may increase false positives and defaults in disadvantaged groups (Hardt et al., 2016); *separation* equalizes error rates (i.e. false negatives), but can still lead to unequal approval rates that reinforce systemic inequalities (Kozodoi et al. 2022); and *sufficiency* equalizes true positive rates across groups, but may allow for disparities in other errors (Liu et al., 2018). These frameworks can then be assessed using statistical metrics like demographic parity, equalized odds, or treatment equality. Some organizations also consider outcome-level measures, such as equal loss ratios (Kelly & Mirpourian, 2021). These conceptual frameworks also influence how fairness is pursued throughout the machine learning pipeline.

Fairness can be pursued at different stages of the ML pipeline: in pre-processing, in-processing, and post-processing (Caton & Haas, 2024; Jui & Rivas, 2024; Mehrabi et al., 2019; Pessach & Shmueli, 2022). Pre-processing may include ensuring that data is balanced across different demographic groups by ensuring each group is equally represented in the dataset or reweighting to account for imbalance (Kozodoi et al., 2022). In-processing methods modify ML algorithms during the training process, such as by optimizing for equalized odds or incorporating constraints in the objective function to ensure a "fair" outcome. Post-processing adjusts model outputs to ensure fairness and enforce fairness criteria, such as by applying calibration, reweighting schemes, or decision threshold adjustments to align outcomes across demographic groups. (Hardt et al., 2016; Jui & Rivas, 2024). This stage is where equal loss ratios across certain groups can be incorporated.

Taken together, fairness in ML-based credit scoring can involve aligning conceptual frameworks with specific statistical metrics and implementing approaches at different stages of the ML pipeline. Each of these reveal decisions that developers must make and the priorities within them. They reveal the value-laden components of fairness. In short, designing for fairness in ML is not simply a technical exercise, but a process shaped by normative judgments about which kinds of outcomes matter, for whom, and why.

6.2.3. Critiques of common fairness approaches in machine learning

Quantitative approaches to operationalizing fairness are limited. First, while conceptual frameworks and statistical measures can serve as useful diagnostic tools, they are narrow in what they capture and can obscure deeper questions about justice, harm, and context. For example, focusing on independence (equal loan approval rates across groups) can miss disproportionate harms in denials, such as higher false rejection rates for marginalized groups. This underscores how different conceptualizations of fairness shape how accuracy is evaluated and what outcomes are prioritized. Similarly, demographic parity may be relatively easy to implement, but can conflict with other fairness criteria, such as individual merit or justice. Second, statistical criteria approaches can conflict with each other, making it impossible from a technical perspective to satisfy all at once (Corbett-Davies et al., 2023). When base rates differ across demographic groups – as they commonly do – it is mathematically impossible for a model to simultaneously satisfy all fairness metrics, such as calibration and equal error rates (Kleinberg et al., 2016). Furthermore, adding different fairness constraints can mean tradeoffs with the overall performance accuracy of the model (Caton & Haas, 2024; Haas, 2019). Developers therefore are forced to choose which fairness goals to prioritize, and within that, what kind of accuracy will be pursued by the algorithm.

While fairness frameworks and metrics offer valuable starting points, they frequently obscure normative questions: Whose outcomes matter most? What kinds of errors are considered tolerable? And who decides what fairness means in practice? These metrics embed assumptions about acceptable disparities and trade-offs, often without making those assumptions explicit. They can also fail to engage with the structural conditions that shape disparities in the first place, thereby not capturing dimensions of fairness that disciplines such as law, ethics, and the social sciences emphasize. Fairness framed through technical accuracy only offers a partial explanation for algorithmic harms, while taking on normative weights in domains like credit scoring (Burrell, 2024).

A focus on quantitative approaches to solving for fairness in ML is not only limited, but can result in unintended consequences. Fairness approaches in ML practice tend to justify solutions that employ surveillance and classification, while masking "broken social systems" that create a "patina of legitimacy" while perpetuating inequity (Pasquale, 2019). In addition, technical fairness approaches distract from institutional structures that produce inequity, making it harder to challenge upstream decision-making (Burrell, 2024).

A prominent example of these tensions is the COMPAS algorithm, used in U.S. courts to assess a defendant's risk of reoffending. ProPublica found that Black defendants were nearly twice as likely to be incorrectly labeled as high risk for reoffending, compared to similar white defendants (highlighting racial disparities in error rates, invoking the conceptual framework of *separation*) (Angwin et al., 2016). NorthPointe, the company that developed the COMPAS tool, defended its fairness by arguing that it was equally predictive for both Black and white defendants, which is an approach aligned with the fairness criterion of *sufficiency* (i.e., equal overall predictive accuracy across groups) (Angwin & Larson, 2016). This example highlights how different fairness frameworks and statistical definitions can lead to contradictory conclusions about fairness – and how developers can use technical reasoning to sidestep deeper questions of justice and equity.

This chapter fills an important gap. While a growing body of work examines technical definitions and critiques of fairness, with some studies exploring fairness considerations implemented and their issues in private (Holstein et al., 2019) and public domains (Veale et al., 2018), there is limited research examining how fairness is understood by those building or deploying these systems, including in financial domains. Understanding how fairness is interpreted, including underlying beliefs, is critical to unpack and move beyond practices of principlism that often rely on abstract, universal notions like fairness that carry unexamined presumptions and superficial solutions to moral problems (Clouser & Gert, 1990). This chapter presents findings to push beyond theoretical debates to critically examine how fairness is operationalized in the real-world context of fintechs leveraging "AI for Good". This empirical lens reveals the tensions, contradictions, and power dynamics embedded in these practices and ensuing gendered implications.

6.3. Methods

This chapter is a continuation of the first chapter with findings drawn from the 25 semi-structured interviews with corporate leaders, data scientists, and investors at fintechs developing and managing

ML-based alternative lending apps in LMICs. The detailed methods for the interviews are outlined in the methodology chapter. This chapter builds from a subsection of the interviews focused on the concept of fairness, which included questions on how fintechs define and operationalize fairness in the context of ML-based credit assessment and loan disbursement. In analyzing data specific to how fairness is understood and conceptualized, I utilized an inductive approach using a grounded theory method. This allowed for themes to emerge and nuance to be captured in how fairness is understood and operationalized. I drew on reflexive thematic analysis, subsequently reviewing and refining themes to identify patterns and core elements in the data (Braun & Clarke, 2006). After familiarization with the data round, I developed an initial generation of codes related to AI fairness. I developed and reviewed themes from the codes before subsequently refining my codebook and writing.

My positionality as a researcher and prior development practitioner led to some nuanced and intriguing interactions regarding fairness. I aimed to ask open-ended neutral questions about the concept, while knowing that the concept carries baggage in the context of machine learning. Some interviewees asked for clarification on what I meant by fairness, others would get uncomfortable or hesitate in responding. I sent all questions to interviewees in advance, yet for several interviewees this was the most tense part of the interview. I felt this stemmed from the fact that it is a very difficult question without clear cut answers and interviewees do not want to be seen as unfair or discriminatory. When some interviewees asked for clarification on what I meant by fairness I tried to respond by saying I wanted to hear whatever they thought about it. Having conducted research on fairness in ML tools before, I did have perspectives on this, but did not want to guide or skew what they shared. This allowed for a variety of different answers and perspectives.

For the analysis in this chapter, I draw on my theoretical framework, with a focus on particular concepts. First, I use Haraway's situated knowledges, which posits that all knowledge is produced from specific, socially, and historically situated standpoints (Haraway, 1988). This lens allows me to examine epistemological positions on fairness and ensuing design choices. I incorporate Social Shaping of

Technology (SST) to analyze how organizational and industry-wide structural pressures (e.g., profitability imperatives) influence fintechs and their employees in constructing fairness perspectives and approaches (MacKenzie & Wajcman, 1999). Finally, I utilize technofeminism to examine gendered implications of fairness definitions and approaches, including how particular fairness frameworks may inadvertently reinforce gender biases (Wajcman, 2006). This is complemented by Benjamin's concept of do-gooderness, which critiques the framing of technologies as solutions to social problems without addressing their deeper, often hidden, implications. This lens allows me to further critique and analyze complexities and pitfalls in how fintechs define fairness (Benjamin, 2019b).

6.4. Findings

6.4.1. Fintechs define fairness in different ways

Fairness can be defined in different ways in the context of ML-based credit assessment. None of the fintechs interviewed desire to be unfair or unfairly discriminatory, and most (with the exception of one) insisted they are fair. However, what fairness means is not always clear, and there is no single definition or approach to achieving fairness across gender or other demographic groups. Lucy, who conducts research and works as a funder of fintechs in this space, shared: "When we talk to digital lenders… They say yes, absolutely it's fair. And then they lack an understanding or an ability to articulate what their definition of fairness is that they're measuring against." The lack of understanding around what fairness means is something that fintechs, as well as researchers in the space share. Jack asserted: "We all agree the question isn't do we want to be fair? The question is, in practice, how does that work given what the tradeoffs are. And so I think a lot of people care about it. A lot of people are researching it. I don't know if we have good answers." Similarly, Sean – a data scientist in a leadership role at his fintech – spoke candidly about their own attempts at grappling with the complicated question of fairness:

> "All the great questions you have listed, that one is by far the toughest one. Because the truth is, I don't have an answer… Would I love to have that? Would I love to be able to say, you know, this

> thing is totally fair and this is why and this is how and it's all good? Obviously I don't have. Why? Well, first we need to find out what fairness is. So you have a metric, right? Or you could do something more like you know qualitative and ask people and whatnot, that's a very expensive process. So I would rather have a quantitative one where I have a formula. Maybe that's just my background. But I would like to say, no, it's fair, that's because this number here is 7.3 and that's fair. But that's not so easy to get… So the only thing I can do, and it's weak and I know it's weak… is just basic sanity checks. So we just sample the data and human beings manually look at things and decide, does that look fair? And I know how terrible this sounds. But I'm being honest about where it is because I think it's a really, really tricky problem."

Sean highlights how there is a subjectivity to considering whether an output is fair in algorithmic-facilitated lending. Similarly, Aditi noted that, "fairness has different implications in different cultures." This recognition that fairness can mean different things in different contexts adds to uncertainty by fintechs.

While there is an uncertainty about what fairness means and lack of clear alignment across fintechs in the algorithmic-lending space, some key trends emerge related to fairness. In particular, fairness discussions focused on what the model produces (the outcome), as well as how that result is arrived at (the process). Findings walk through these trends, while also exploring their implications.

6.4.2. Process: Fairness is achieved by…

Most fintechs interviewed are not considering fairness in the design and implementation of their ML models or note that they are achieving "fairness through unawareness". From a gender perspective, this means that if gender is not considered by the model, it is fair from a gender perspective. This is tied to the logic of "leaving it to the machine", whereby approaches to being fair mean not having humans get in the way of the (objective) technology.

In some cases, interviewees noted that they did not see fairness as a priority, nor did they consider how to operationalize it. Gary responded when asked about fairness: "It is an interesting question… because by and large, we are relying on our credit risk engine and what our machine learning actually presents to us." On the surface, this appears as if they lack a fairness approach, however it is in line with achieving fairness by being unaware (i.e. "fairness through unawareness"), which is aligned with the belief in the objectivity of ML and AI. Dante put it bluntly:

> "The short answer is that we don't [consider fairness]. Because we don't bring bias into the model. We don't know the gender, we don't know the ethnicity. I smile at your question because it's kind of a US kind of question… Believe me, nobody cares, outside of the US or perhaps the UK, if the model is fair or not."

This response highlights how perceptions of fairness can often be linked to social justice – and that algorithmic-facilitated lending is not the appropriate place to advance social justice. Rather, algorithmic-facilitated lending is seen as apolitical and objective.

Others more explicitly attached fairness to approaches of being "blind" to gender. While not focused on social justice, these perspectives support the belief that algorithmic-facilitated lending is objective and can cut through unfair human-caused bias and discrimination. The overarching sense is that ML makes lending more fair and responsible by avoiding bias that humans have through using machines that detect patterns and make decisions based on a set of rules. Therefore, the act of using ML is what makes the tool fair. From this perspective, fairness is achieved not necessarily through a careful design of outcomes or constraints, but through the very act of delegating decisions to the machine. As Michael put it, "As long as the AI model is trained the right way like... it actually is a huge net positive… your demographic data should be totally excluded." If the machine is developed well and does not factor in certain protected characteristics, fairness is then an outcome.

Beyond being gender "blind", there is another approach taken in line with "fairness through unawareness". This is the approach several fintechs take when entering markets in which they proactively give out loans randomly as part of the process in building models. Here, giving out loans randomly and seeing who repays allows fintechs to identify predictive variables for repayment and creditworthiness. Megan explained, "If we have a budget accepting, you know 10,000 users a month, we accept 10,000 users a month. And that helps us get initial data on how to start grouping people into high risk, low risk, medium risk." They can also then refine models over time based on repayment trends. As discussed in

Chapter 5, this method allows for the machine to identify variables correlated with who repays loans as opposed to data scientists or fintech leaders relying on their own judgment in identifying variables for the model. It can still involve some aspects of choice, in regards to what to incorporate into the model that is applied to new users, but is a method that interviewees discussed as ensuring and supporting fairness.

Some interviewees noted that the tool is fair because it is accessible to people, especially marginalized groups. As Rakesh put it, "The way we look at fairness is that this is a system that is open to anyone… This is a free to use app by anyone who wants to access bank loans. There are no further limitations on this beyond just having their mobile phone [and] being able to download an app." Here, the tool is fair because it is available for anyone to hypothetically be able to use, while owning a mobile phone and being able to download apps is seen as basic and easy. This reflects a fairness claim rooted in how the system is designed and set up (i.e. fairness as the access and rules are the same for everyone so participation is technically possible), but does not necessarily include who actually benefits from the tool or whether outcomes are equitable. This was echoed by Suraj, "I don't want to call it fair, but not because it's unfair, but I just, I think that doesn't apply. Like it's an ML based model, so you can scale to everyone." In this, fairness as a concept doesn't apply as the tool is accessible to people and can reach scale. These perspectives are also in line with "leaving it to the machine" where fairness does not need to be considered in the model itself.

6.4.3. Outcome: Fairness is achieved when… ?

Interviewees also discussed "fair" outcomes, sharing perspectives about what was achieved if a model was operating fairly. In sharing perspectives of fairness as outcome-based, fairness is judged based on what the model achieves, versus how it gets there. This outcome-based view of fairness puts an emphasis on effects of the model in the real-world, such as increasing financial access to people who are underserved. In this, fairness is validated by the result.

6.4.3.1. Fairness is achieved when securing social impact

Various interviewees see fairness as being achieved simply by doing good. Here, tools are fair because they are advancing financial inclusion for people who would otherwise be financially excluded. Fairness is achieved by having a social mission that allows people to enter financial folds, and/or because the ML tools are better than the status quo. Suraj asserted: "If this product is focused only on one segment, like every loan you sell is a fair loan. You know what I'm saying? …. Every loan you make is very impactful because you're giving loans to people who are pulling hand carts or electricians, day laborers. Right?" Various interviewees echoed this sentiment noting that financial inclusion is the goal of these tools and they are fair because they are enhancing financial inclusion and reaching underbanked people. As Jack noted, "If financial inclusion is the goal, it's fair because it's reaching people who are underbanked. So there's kind of like a socioeconomic piece to the fairness explanation." Similarly, some interviews discussed fairness as being better than the status quo. This is a common sentiment of being fair because it is better than the status quo and avoids some of the historical biases in finance that have inhibited women from accessing credit.

Other interviewees take this further and see fairness not just as doing good, but doing as much good or securing as much benefit as possible (i.e. benefit maximizing). Here, inclusion is prized: More people accessing loans is the most socially impactful approach and therefore the most fair. This sentiment of fairness as inclusion is reflected in the aforementioned quote from Rakesh, whereby the tool is fair as anyone can hypothetically access the free app. This approach prioritizes inclusion over considerations of equity. It also does not consider how many women versus men benefit, rather considers the number of people overall. Some fintechs track these numbers to understand their reach, but don't have particular goals related to loan distribution across gender and other demographics. Nicholas shared: "I don't believe that it necessarily has to be extremely complex… just to see how these customers fare compared to other segments of customers." The goal is simply tracking to understand the numbers.

No interviewees discussed fairness in terms of parity – whereby both men and women have equal probability of getting a good credit score and access to a loan. In fact, this is generally looked at negatively by interviewees who do not see the role of fintechs being to correct for inequalities that exist in society. Megan explained that tracking approvals by gender is:

> "A very black and white way to look at it... You can't say that [if an] approval rate of women versus men is higher in India that our models are negatively biased against men. That is not a really good measure… You can't really just look at all approval rates and [say] they have to be the same."

From a lender and business perspective, this approach is logical, as loans are tailored based on various factors, many of which are shaped by structural inequalities.

6.4.3.2. Fairness is achieved when the tool is accurate

Several interviewees had a more detailed and technical response around when fairness is achieved. They discussed that fairness is achieved by the ML tool when it accurately predicts creditworthiness across different protected groups. John, who is an engineer trained in ML and leader of a large fintech, noted that he thinks of bias from the technical definition in regards to ML. In this, if you're consistently wrong in the same direction that would be biased, such as being consistently wrong about women. But if they are correct about people (even if they are more often giving loans to men than women) it would not be biased. He continued that women in the countries where they work have less money on average than men. So one would want to give them a lower credit score, and this lower credit score would be accurate in regards to whether one should get a loan (even though it reflects inequality). Relatedly, giving out loans to people who could not repay them can have unintended consequences for that individual so it is important to ensure the person can repay the loan. John noted that this is the reason why lending is not a good way, therefore, to correct for social inequality.

From a lender's perspective, some degree of bias is seen as inevitable, and the primary goal becomes delivering accurate credit decisions – offering the "right" amount of credit to individuals likely to repay. While this might seem like a straightforward objective, complications arise: what exactly should be measured to determine if a tool is accurate? And what outcome should the model be optimized for – minimizing defaults, maximizing profit, or expanding access? These questions reveal that even the pursuit of accuracy involves value-laden decisions about goals, trade-offs, and priorities. Megan described their process of comparing credit margins and optimizing across risk groups:

> "I don't try to optimize that much against female and male. I look across multiple groups, I look across risk groups… credit score deciles. I look across [credit] limit groups, like for borrowers that have a zero to $100, versus $100 to $200, versus $300 to $500… Essentially the main metric that I look at is credit margin. What is my credit margin across customer risk segments? And credit margin is literally just, how much revenue did you charge minus default rates, over revenue… If you look at approval rate differences, that is not a good enough sign. You have to double click into what is the loss ratio by segments. Loss ratio is essentially defaults over revenue… If you're receiving more in revenue given a group's level of risk, that is unfair. And either you have to find a way to adjust the interest rate down of that protected class or increase the interest rate of your other groups to get to that uniform loss ratio."

This type of approach considers the revenue that customers provide and aligns financial margins across different risk groups, regardless of demographics. Here, a customer paying late may not be considered a negative outcome, as long as they ultimately repay the loan and the financial return is acceptable. This represents a "post-processing" approach (i.e. adjusting model outputs to ensure a fairness criterion) and reflects a quantitative, revenue-centered notion of fairness rooted in the lender's business objectives. This is not an uncommon approach in lending. Indeed, Megan shared that this approach was taken after careful consideration and reading reports from the industry on the topic. This approach represents a more nuanced interpretation of accuracy, which goes beyond simply predicting repayment to account for the timing and profitability of that repayment. This illustrates how definitions of fairness – and accuracy – can vary and depend on the perspective of those developing and managing the algorithms.

Notably, Megan was the only interviewee that explicitly described using a fairness metric and statistical approach – equalized loss ratios across risk groups – while others either did not mention fairness measurement or gave no indication that they conducted audits or evaluation. The lack of standardized testing and transparency across most interviewees reflects an overall immaturity around how fairness is considered and operationalized in the ML-based lending space in LMICs. Despite widespread perceptions that their models are fair, actual implementation and measurement appears largely ad hoc or absent altogether.

While others did not have or share a particular fairness metric, this type of post-processing method shared by Megan may be present in other fintechs. Lucy, who works at an organization that funds and conducts research with fintechs, noted an interesting phenomenon they had found:

> "For nearly all of the institutions we've worked with, when we look at the actual portfolio, it looks really balanced… like credit scores are the same among men and women in the portfolio. But then when we look at applicants, we see that women have a higher credit score or a higher rating in the internal assessment of creditworthiness. But they're not being given credit for other reasons… That's the one we find over and over again."

This illustrates an interesting phenomenon interviewees report observing: women are getting denied loans more often who have higher scores compared to men who are denied loans. Lucy guessed the issue was linked to unconscious bias (e.g., if the lender had some human intervention in the design) or if women had more missing data. However, post-processing fairness approaches that ensure equalized loss ratios could also provide an explanation.

6.4.4. Governance and bias mitigation as add-ons

Alongside process- and outcome-based understandings of fairness, a handful of interviewees raised issues of organizational governance, explainability, or bias mitigation. These did not form part of their definitions of fairness, but rather surfaced as add-ons in the discussion. Some admitted uncertainty about what bias even meant in practice; others referred to explainability tools or internal governance structures. These scattered references suggest that responsibility is recognized, but only loosely and inconsistently tied to fairness, reinforcing the fragmented nature of how fintechs approach the issue.

A couple interviewees discussed governance approaches in the organization that are set up to support responsibility considerations. Megan, who works at a fintech with global operations that is one of the rarer female-led fintechs in the space, shared that ethics is a common discussion. She discussed being part of the organization's Ethics Committee, which is the only company interviewed that had such an entity: "I'm really passionate about the subject [of fairness] and I was part of the Ethics Committee at [fintech company] where we did a lot of deep dives into this [question of fairness and assessing it]." This was the only interviewee and fintech that discussed broader organizational structures to support ethical discussions and considerations.

Processes to assess for and mitigate bias, as well as consider other responsibility questions, is a constant work in progress. Several interviewees noted they were curious if I had ideas for them. In discussing how to mitigate bias, Sean said, “That is a really tough one… I would love if you had some ideas here on what we can do.” In general, there is a spectrum of approaches with some more rudimentary than others. Sean continued: “Let’s pick two people at random. Let’s see what the scores are. Then let’s just look at it with our human eyes. You know, what do we think about this individual or this group?” He is a leader at a fintech company that is newer to the credit assessment space and they are still refining their approach. Rohan, on the other hand, shared that they have “fairly detailed” senior-level reviews to ensure there aren’t unintended biases.

Lucy, who works at an organization that funds fintechs and provides research support, shared about a bias scorecard they created for lenders. The scorecard recognizes that bias and fairness can be defined in different ways, so it includes several metrics to provide a more holistic gender picture. However, it hasn’t gotten picked up as much as they would like. This is in line with my interview data in which participants did not discuss measurement or auditing approaches (with the exception of one).

Another common factor that interviewees brought up as part of efforts to support responsibility are explainability and transparency. Rohan discussed that his fintech wants to be able to explain a decision for two reasons: first to ensure they know why a borrower is considered risky (or not), and secondly, to leverage those insights to help people build their own financial literacy. Both of these things helped the organization build credibility.

However, as machine learning is considered a “black box”, explainability can be difficult and the extent of the explainability depends on the particular model. For example, more basic decision tree models are more transparent as decisions can be visualized as a tree structure, making it easier to understand how the model arrived at a particular outcome. On the other hand, more complex models like neural networks and ensemble models (e.g., random forests, gradient boosting) that fintechs often use can have challenges with transparency and explainability as there are more intricate structures and interactions among features. They can produce high-quality predictions, but decision-making processes are more difficult to interpret. This illustrates tensions between more technically complex and powerful models with transparency. Generally there is a lack of transparency provided to users about how the credit scores and lending decisions were arrived at. This is partly linked to fintechs not fully understanding in many cases, as well as keeping them purposefully opaque.

One interviewee noted how the "black box" nature of algorithms and lack of regulation can complicate understandings of fairness. Kamal reflected on how the Indian regulatory system is different than the US. He noted, "For good or for bad, we don't have FMLA or whatever you guys call it over there right? … We do not have this here. So obviously we want to be equitable, but typically models are black boxes… What do you do? So we don't have that issue right now." The lack of regulation in this case meant that they did not need to be concerned about discrimination for the time being.

6.4.5. The role of investors and market incentives in shaping fairness

Investors play a critical role in shaping priorities and informing approaches of fintechs. Several interviewees discussed pressure from investors as a reason they had to update an approach or focus on easier-to-reach customer segments. Hannah, who works as a researcher at an organization that supports and invests in fintechs, shared:

> "We're pushing them a bit on fairness and responsible AI and they would say, well you know we have to keep in mind the incentives we have from our investors. We need to hit certain growth targets. We have certain KPIs that are part of our covenants or a part of our agreements with investors and that might lead us to prioritize low hanging fruit, which are customers who are perhaps already included or have more robust data trails. But we would need more resources to build out more checks for fairness or even to reach more excluded customers to build models for them."

Many of the fintechs providing ML-based alternative lending tools are startups or small businesses operating in a tough industry. In the last couple of years, as a result of various macro-level factors including high interest rates globally in recent years, venture capital fell drastically leading to a “fintech winter”. Indeed, in 2023 venture capital in fintechs decreased 42% (S&P Global, 2024). In many cases, the fintechs are trying to keep the doors open and are pursuing profitability. Gary explained:

> “We are a startup… and many of the players out there that you’re speaking to are startups. So we’re running the businesses and ultimately the business model we need to manage for profitability and for lending business. It’s really managing your loss rate, especially with [the] cost of funds going up. It makes this a very tough business to actually run. I think everybody that you’re speaking to generally is loss making as a business.”

Limited resources (time, money, people) can impede more robust responsibility approaches, while pressures to be profitable can push fintechs towards lower hanging fruit. At the same time, business pressures push fintechs toward efficiency and profitability, often shaping how their models are developed and evaluated. As discussed in Chapter 5, several interviewees noted that the models are predicting creditworthiness while optimizing for lifetime value (LTV), a common metric used in lending, that estimates revenue over time minus acquisition and servicing costs.

Kumar shared tensions between priorities of making money and achieving fairness for different types of groups, including women. He said, “[The fintech is] trying to make money so you can’t, you know, push fairness when you look at income bands… I don’t think the models are that fair… Maybe housewives would probably be in a less fair situation to get a higher ticket size because, you’re not taking gender [into account but in an] employment setting, [they are] probably homemakers.” This honest assessment applied understandings of parity in considering fairness, while exposing how priority on profits impacts fairness in practice.

6.5. Discussion

There are different understandings and maturity levels of fairness, with many fintechs lacking a formal or structured approach to incorporating fairness in their models. Fairness is commonly discussed from the perspective of process (i.e. what is a fair *approach* when using ML for credit assessment) and as an outcome (i.e. what is a fair *outcome* from using ML for credit assessment). From a process perspective, the most common approach is "fairness through unawareness," a view that reflects a deep trust in the objectivity of ML technologies. From an outcome perspective, a common response was that fairness is achieved by doing good or benefit maximizing. Others equated achieving fairness by being accurate. Both of these approaches tend to ignore questions of equity or social justice: the former is aspirational but vague, while the latter prompts mathematical evaluations of fairness. Among the few interviewees that conduct mathematical evaluations to assess fairness, those evaluations are oriented around business goals, reflecting profit priorities. Alongside these perspectives, some also raised responsibility considerations and bias mitigation, but in fragmented ways: uncertainty about what counts as bias, reliance on explainability tools, or reference to organizational governance approaches. These scattered practices signal that while responsibility is part of the conversation, it remains loosely connected to how fairness is defined. Taken together, these findings illustrate a fragmented and business-oriented approach, ranging from reliance on technical objectivity to profit-driven evaluations, with responsibility and bias mitigation treated as add-ons rather than as part of a coherent, equity-centered framework.

The following discussion is organized around five interrelated themes that emerge from the data. These themes reflect not only how fairness is understood and applied by fintech actors, but also how those understandings are shaped by assumptions about ML objectivity, institutional goals, and incentive structures. First, I discuss how fairness is often declared rhetorically but not embedded meaningfully in models. Second, I explore how fairness is subjective and largely informal, with few mechanisms for measurement or enforcement. Third, I argue that even seemingly neutral concepts like accuracy are value-laden and introduce the concept of *situated priorities*. Fourth, I show how institutional logics within fintechs (i.e. profit) dictate understandings and approaches for fairness, with gender implications. Finally, I consider the broader limitations of relying on market-based actors to deliver equitable outcomes in the absence of regulatory pressure. Together, these points illustrate how fairness is often invoked rhetorically or instrumentally, and how these framings contribute to the reproduction of gendered inequalities in lending.

6.5.1 Fairness by declaration, not by deliberate design

A core theme that emerged in the data is the widespread invocation of fairness as a declared value – rather than a principle embedded in design decisions or model evaluation. Fintech actors often describe their systems as fair because they use ML, expand access, or serve a social purpose. However, they don't walk the talk, as they rarely articulate how fairness is actively defined, implemented, or measured. "Fairness through unawareness" is a common example of this rhetorical framing – and is fraught. The idea that machines, when left alone will be fair, ignores the ways that data is rife with biases. Indeed, data reflects limiting stereotypes and norms about people, while datasets can under-represent certain groups leading to lower performance (Criado-Perez, 2021; D'Ignazio & Klein, 2023). It also ignores how features and proxies used in models carry gendered norms and reflect structural inequality (D'Ignazio & Klein, 2023).

The perspective of "fairness through unawareness" is tied to the belief that the act of using ML makes a system fair. This perspective – in line with the Silicon Valley logic of "leaving it to the machine" – suggests fairness is not achieved through explicit constraints or auditing, but through the act of delegating decisions to the algorithm itself. Fairness becomes framed as a property of a method and as an automatic outcome of using ML: If data is the truth and the machine is objective, the results must be fair. The danger in this view is that it discourages critically examining outcomes or considering structural inequalities. It overlooks the ways that data or optimization goals (e.g., profitability) can inadvertently penalize some groups. The framing urges a hands-off approach to accountability in which gender differences in algorithmic lending can be hidden and/or legitimized.

A second common rhetorical framing of fairness centers on outcomes, particularly the idea that simply expanding access or delivering positive results constitutes fairness. Fairness as benefit maximizing or doing good echoes a consequentialist logic, whereby fairness of a system is judged by its outcomes (e.g., how many people receive loans) rather than how decisions are made. This logic often appears alongside procedural notions of fairness (e.g., fairness through unawareness and being "blind" to gender). The blend reinforces the belief that if a model applies the same rules to everyone and helps people, it must be fair. Still, it obscures structural inequalities and avoids scrutiny of who is benefiting or not, while justifying doing so under the umbrella of being for good.

This echoes Benjamin's concept of do-gooderness, which critiques the framing of technologies as solutions to social problems while ignoring deeper, hidden implications, including through using narrow definitions and operationalization of fairness. Benjamin argues that algorithmic racism often arises from this moral cover, which overlooks power hierarchies under simple, mathematical framings of fairness. In algorithmic lending in LMICs, declarations of fairness grounded in claims of doing good masks unintended consequences. It becomes a blindfold, even to the developers and managers of the fintechs themselves. The result is an alluring – yet selective – fiction about the potential of "AI for Good" to lift marginalized communities while leaving gendered and structural inequities unexamined.

6.5.2 Fairness is subjective and unaccountable

While fairness is often invoked, interviewees rarely described it as operationalized in systematic ways. The disconnect between rhetoric and practice stands in contrast to the academic literature on ML, which offers a proliferation of fairness metrics (e.g., demographic parity, equalized odds) and strategies for intervention across different stages of model development (e.g., in-processing, post-processing). Despite this, many fintechs either aren't engaging with these fairness metrics and methods or are doing so in fragmented, inconsistent ways. Some fintechs did have approaches, like evaluating and optimizing for equalized loss ratios in post-processing, but most do not appear to systematically consider, measure, or audit fairness in their models. This highlights a significant, related disconnect between perceived and demonstrated fairness. Many organizations operate under the belief that their models are fair, without systematically testing for or measuring fairness across groups. This is also reflected in the scattered responsibility and bias mitigation approaches across fintechs that are ad-hoc and add-ons.

It's not that fairness claims are disingenuous; rather, they reflect a common belief that fairness has already been achieved or is inherently embedded through using "objective" ML and/or pursuit of social impact. Drawing on Haraway's concept of situated knowledges, these assumptions can be understood as epistemological positions that present fairness as a universal, neutral standard; yet in reality, it is constructed from specific, power-laden standpoints. By framing fairness as something that has already been achieved, they render external evaluation, formal metrics, or auditing unnecessary. As a result, fairness functions more as an implicit belief emerging from these logics, rather than something to explicitly consider in the design of the model or be empirically tested and audited for. Fairness thereby becomes subjective and unaccountable: something that doesn't need to be measured, because it is a presumed outcome of machine learning objectivity and good intentions.

Compounding these logics are other barriers that limit uptake of fairness practices. Following the Social Shaping of Technology framework to examine industry-wide influences, the sheer number of fairness definitions, metrics, and techniques (without clear consensus or regulatory guidance) may be part of the problem, contributing to confusion, inaction, or ad hoc adoption. Even among those who recognize that considering fairness is important, they often lack resources or incentives to build robust approaches, operating under tight constraints reflecting broader structural pressures in the fintech sector. In addition, they operate in various markets, which tend to have immature AI regulatory landscapes. These conditions reinforce treating fairness as an assumed quality versus something to rigorously design for and pursue. The result: systematic approaches to fairness and responsibility remain the exception rather than the norm.

The one organization that shared a more robust fairness and responsibility approach reveals the critical role of leadership. By establishing an Ethics Committee and prioritizing fairness from the top, leadership legitimized it as a meaningful concern. They also encouraged critical staff engagement and allocated resources to support it. Interestingly, this organization is one of the few fintechs in the interview sample that has high levels of female leadership.

6.5.3 Accuracy is in the eye of the developer

Several fintechs connected being fair to being accurate. However, accuracy in credit assessment is not a fixed concept: it can be defined or pursued in various ways and depends on what outcome the model is designed to predict. For example, a fintech company might be defining accuracy in terms of correctly predicting whether someone will default at all, whether they will pay on time, or how profitable they will be as a customer. Each of these reflects different situated knowledges and institutional logics, emphasizing how what counts as accurate is not neutral, but shaped from specific standpoints and organizational pressures.

Framing fairness as accuracy aligns with dominant ML practices where fairness is often treated as a mathematical optimization problem, despite there being different ways to consider fairness in social science and philosophical disciplines (Mulligan et al., 2019). Furthermore, creditworthiness assessments can become self-fulfilling prophecies, complicating the idea of accuracy. For instance, if someone is deemed a riskier borrower, their loan offer and repayment conditions are adjusted (e.g., higher interest rates). If they default, the algorithm could be seen as accurate for determining them to be a riskier borrower. But with better loan terms, they might have successfully repaid. As illustrated by (Citron & Pasquale, 2014):

> "[Credit] scores can become self-fulfilling prophecies, creating the financial distress they claim merely to indicate. The act of designating someone as a likely credit risk (or bad hire, or reckless driver) raises the cost of future financing (or work, or insurance rates), increasing the likelihood of eventual insolvency or un-employability. When scoring systems have the potential to take a life of their own, contributing to or creating the situation they claim merely to predict, it becomes a normative matter, requiring moral justification and rationale."

In economic and industry discourse, discrimination and bias is not necessarily bad in lending. Actually – and tied to American banking logics – lenders see it as rational, absorbed into the definition of accuracy, and inherently part of the lending process. In lending, discrimination is framed as identifying “good” versus “bad” risks, even if these judgments perpetuate social inequities. The 1974 Equal Credit Opportunity Act in the US did not ban discrimination, but rather clarified that discrimination must be unfairly discriminatory. Here, women denied credit at higher rates than men could be seen as discrimination; however this is not at the fault of lenders, but rather labor markets that reflected social and economic disparities in American society (Lauer, 2017). It would therefore be acceptable. Researchers have shown that biases against female loan applicants can be seen as rational responses from lenders; however, it has consequences, including perpetuating gender segregation in lending and becoming self fulfilling as unequal access further entrenches differences (Buvinic & Gokhroo, 2023).

Interviewees reflected this perspective, arguing that gender differences in lending is rational, as being fair means predicting creditworthiness accurately, not rectifying societal imbalances. While fintechs don’t intend to cause harm, they also don't see themselves as responsible for fixing structural inequalities. This illustrates how fairness framed through accuracy can overlook social science definitions, which emphasize fairness through a lens of justice and addressing systemic inequalities. Equating fairness with accuracy in ML may appear objective, but only offers a partial explanation for algorithmic harms and can obscure broader social impacts (Burrell, 2024).

The varied ways that fintechs define and pursue accuracy reveal that it, too, is a subjective and value-laden concept – shaped by specific standpoints, organizational pressures, and industry-wide perspectives. This intersection forms what I term, *situated priorities*. This concept addresses a gap in existing theories: frameworks like SST focus on broader social and structural pressures while situated knowledges emphasizes epistemological positioning; yet, neither fully examines how individual actors negotiate and prioritize specific goals within those structures that become embedded in the technology. Situated priorities includes several key propositions: (1) priorities are situated, constructed through a combination of epistemological positioning and institutional logics; (2) priorities are dynamic and negotiated within decision-making and operational processes; and (3) priorities shape different technological outcomes by defining what matters most and embedding those priorities in technology-related decisions. Situated priorities offers both a framing and theoretical lens for understanding how actors' decisions are shaped by what they consider important or valuable – priorities that are formed through their institutional contexts and situated perspectives. It enables analysis of how principles and values are navigated, defined and implemented in practice, and how they evolve over time. It also helps trace how abstract principles (e.g., fairness) are interpreted, negotiated, and translated into particular goals, leading to diverse technological outcomes. This reveals how values are not simply applied, but actively shaped through organizational dynamics, practical constraints, and sociotechnical decision-making.

6.5.5 Fairness is shaped by institutional incentives and priorities

While fairness is often discussed as a technical or ethical question, in practice it is deeply shaped by institutional incentives and market logics. Indeed, technology is not only constituted through human agency, but through the strong influences of organizations and institutions, resulting in outcomes of technology that are a result of this complex interflow (Fountain, 2001).

Market logics and priorities impact what ML algorithms are optimized for (e.g., profitability, scalability, risk mitigation), as well as how fairness and accuracy is assessed (or not). Several interviewees described models being optimized for business objectives like LTV, reflecting profit priorities. Fairness constraints were not mentioned during model development. While several interviewees discussed balancing loss ratios, one organization described their post-processing equalized (or uniform) loss ratio approach.

6.5.5.1 Equalized loss ratios

Under equalized loss ratios, fairness is defined in terms of financial parity: ensuring that no group is disproportionately risky relative to the revenue they generate. If one customer risk segment has a higher loss ratio than others, the model's outputs are adjusted by rebalancing approvals or pricing to make loss ratios more equal across groups. This approach reflects a business-driven fairness metric seen in lending and insurance (Schreiber, 2019), which has been used by other fintechs as well (Kelly & Mirpourian, 2021). It departs from fairness frameworks and metrics in the AI ethics literature, which focus on conceptual frameworks like separation, and metrics such as demographic parity or equalized odds. Fairness in this case is framed from a risk management and financial lens: no group should systematically generate more profit relative to its risk. There is still a choice in terms of what is grouped and measured (e.g., demographic groups, risk groups) and how outcomes are interpreted.

Importantly, this approach doesn't aim to eliminate differences in approval outcomes; rather, it accepts outcome differences as long as resulting margins or loss ratios are uniform. While appealing in industry contexts, this logic raises normative concerns. It may systematically favor groups that generate more money, even if they are less reliable. Also, equal loss ratios can obscure differences in access: a model may equalize loss ratios while still denying disproportionately more loans to historically marginalized groups, or offering them worse terms. Therefore, while compelling in financial spaces, it has limitations related to justice and social equity.

Quantification thereby serves as a moral veil, obscuring the consequences even from those doing the quantifying. In his examination of kidney allocation algorithms, Robinson discussed quantification as "a moral anesthetic", in which numbers seem dispassionate and impartial. Numbers and algorithms help "find a way through the moral conflict" (D. Robinson, 2022). In this case, I argue that quantification is not necessarily serving as a moral anesthetic, as much as a moral veil for certain decisions. This moral veil happens somewhat unknowingly as developers haven't fully grasped the implications and consequences themselves. This is partly a consequence of the misguided and widespread belief that algorithms are technical – as opposed to sociotechnical – systems.

6.5.5.2. Profit as the priority is not new in credit scoring and digital financial inclusion

Profit as a priority is not a new concept in credit scoring, nor in digital financial inclusion. The US credit scoring market evolved alongside capitalist objectives (Lauer, 2017). In a study on the US credit scoring model, Hohnen et al (2021), puts it: "Credit evaluation becomes a calculation of possible profit generation rather than an estimation of a probability of default. A credit score is no longer just a risk assessment, it is a value assessment in a commercial sense" (Hohnen et al., 2021). The focus on profit is echoed in digital financial inclusion in LMICs more broadly. Akolgo (2023) explores how mobile money solutions in Ghana intentionally extract profit from poor people with unintended consequences, such as mounting debt. Similarly, Langley and Leyshon (2022) illustrates how JUMO, an ML-based credit assessment fintech, generates credit scores primarily meant to generate profit and, in doing so, deny people what is actually needed for emancipation. Both studies highlight and connect to a body of postcolonial research that highlights how digital financial inclusion interventions – developed by foreign fintech firms and funded by shareholders largely from the Global North – commodify and monetize poor people under the guise of financial inclusion, without addressing underlying inequalities and economic issues that persist (Akolgo, 2023; Gabor & Brooks, 2017; Jain & Gabor, 2020; Langley & Leyshon, 2022). When fairness is shaped by profit imperatives, its gendered consequences are important to consider, especially in contexts where structural inequalities already limit women's access to finance.

6.5.5.3. Gender implications of fairness under a profit orientation

Applying a technofeminism lens can help to uncover how these profit-oriented framings of fairness have gendered implications linked to the financial behaviors of men and women. In contexts where women are better repayers but take smaller loans (a pattern noted by the vast majority of interviewees and discussed in Chapter 5), approaches like lifetime value and equal loss ratio may fail to reward women's reliability. Larger loan sizes generate higher revenue, which can offset the costs of late payments and defaults. Smaller loans, on the other hand, may incur higher transaction and administrative costs per dollar lent (A. J. Blanco-Oliver et al., 2023). If fintechs optimize for lifetime value or apply equal loss ratios without appropriate safeguards they risk replicating gender bias by penalizing women for generating less revenue.

Men could be overextended if they are allocated too much credit at higher risk thresholds, increasing the likelihood of overborrowing or falling into debt traps.

This would not be the first documented case of ML tools in lending exhibiting a bias due to optimizing for profit: Research on a peer-to-peer lending platform in China found that the introduction of ML to inform interest rates resulted in higher interest rates for women. This was not due to a higher estimated risk or lower determined creditworthiness by the algorithm, but because women had lower price sensitivity and the platform could therefore better optimize revenue by offering women loans at higher interest rates (Chu et al., 2023).

At a higher level, if models optimized for profit are learning that men are more profitable, they may learn to downgrade women altogether. This can happen by identifying women through certain proxies (as discussed in Chapter 5 on how gender can be learned). This pattern of a certain identity being unintentionally and systematically downgraded has occurred in other algorithms; for instance, the previous example shared about Amazon's hiring algorithm downgrading women's resumes (Dastin, 2018). Models can connect optimization goals to certain identities, and then learn and use demographics, like gender and race, through proxies in ways that are not understood and difficult to anticipate. In lending, profitability is not a neutral benchmark. Ignoring how gender norms impact economic behavior exposes models and the fintechs that operate them to risks, including amplifying harmful discrimination and disparate impacts, whereby women are systematically downgraded despite their reliability.

Importantly, as actors begin to recognize that their own perspectives and situated priorities shape technological outcomes, a reflexive moment emerges. This is reflected in the development of the aforementioned Ethics Committee at the fintech where Megan works, as they are trying to grapple head-on with fairness definitions and approaches. This is also reflected by Sean. He acknowledged in our interview that identifying what fairness means is tough and his fintech's approach is influenced by resources available and his own background, before then asking me if I had ideas. These reflexive moments open space to question the neutrality of AI systems and surface tough ethical questions.

6.5.6 Businesses, not "social justice warriors"

Fairness cannot be separated from the priorities of those who build, manage, and fund these systems – in this case, within for-profit and social impact-oriented startups. The perception of fairness as doing as much good for as many people, or benefit maximizing, is tied to business priorities and goals of market expansion, while also linked to priorities for social impact scale found in the development industry. This focus on inclusion for as many people as possible justifies gender differences in lending. Prioritizing inclusion over equity leads to algorithmic tradeoffs and can also mean solidifying power hierarchies and inequities. Meanwhile, measuring and optimizing models against fairness perspectives that are grounded in profit priorities can justify discriminatory practices that uphold – or even worsen – existing gender economic inequalities. The pattern of encoded gender norms, in which the status quo of gender inequality is solidified, is upheld and potentially amplified as optimization focuses on organizational priorities.

These findings underscore how ML-based credit assessment algorithms are normative, reflecting situated priorities. In a study of credit scores in insurance pricing, Barbara Kiviat notes, "Algorithmic prediction is imbued with normative viewpoints – they are viewpoints that suit the goals of the corporation" (Kiviat, 2019). Profit priorities of corporate institutions intersect in algorithms, under guises of machine learning's objectivity. Sadowski reflects on machine learning in insurance, highlighting that truth is judged not on scientific soundness, but on "its pragmatic power for profit-making… which deny an inferior subjective viewpoint and assert a superior objective perspective" (Sadowski, 2025). Indeed, data-driven scoring systems for risk classification are not new in the financial sector and include scientific and technical methods to justify their interests and actions (Fourcade & Healy, 2024). This dynamic reflects Ruha Benjamin's concept of do-gooderness, illustrating how ethical narratives are co-opted by institutional logics focused on scalability, efficiency, and market expansion. In this framing, inclusive and "fair" technology becomes a mirage for reputational legitimacy, even as it potentially projects existing inequalities into the future. Without transparency of models and regulation, “fair” ML systems can entrench feedback loops that reinforce an unjust status quo.

While arguments can be made that it is not the responsibility of fintechs to solve structural inequalities and injustices – they are, after all, businesses not “social justice warriors” – it nonetheless exposes how certain values and priorities inform technology design and management that negate the claim that the technology is objective while also exposing potential unintended consequences. Considering or addressing structural inequalities in the design and management of algorithms is ultimately a choice that must be considered consciously, particularly in the context of for-profit companies seeking to do good.

At the same time, as John pointed out, giving loans to people who cannot or will not repay them (which may result in contexts where “social justice warriors” insist on giving loans equally to men and women or to those who are more marginalized) also has unintended consequences and can result in immense harm. Fairness does not necessarily mean parity or provision of equal loans. Various research has exposed how giving more loans to vulnerable populations (often at higher interest rates) reflects a sort of “reverse redlining” whereby people can be targeted for loans under worse conditions with consequences such as over-indebtedness and debt traps (Garcia et al., 2024). This illustrates the tricky moral and operational balance that fintechs must walk: enhancing inclusion and financial sustainability in a space where both inaction and overcorrection can produce unintended consequences.

Ultimately, fairness in ML is not just a technical challenge for which to optimize, it is a socio-technical problem tied to institutional priorities, value judgments, and power. While statistical fairness frameworks can be useful tools, they obscure deeper questions: Who defines fairness? What harms are considered tolerable? Whose priorities are reflected in and who benefits from these decisions? In for-profit domains such as lending, economic values and priorities bubble to the top. In the absence of regulatory guardrails and accountability mechanisms, fairness risks being defined in ways that serve organizational interests over social equity. Addressing fairness requires grappling with these realities – and being transparent about how fairness is defined, who it serves, and why certain choices are made, while recognizing the limitations of for-profit actors to self-regulate.

Fintechs can take various actions, spurred through reflective moments when actors recognize the influence of their perspectives and situated priorities. First, they should clearly define their fairness approach and be transparent about tradeoffs. Second, transparency can be improved through tools like model cards, which detail fairness frameworks, metrics used, and model-stage interventions. Third, internal and external audits are crucial for identifying disparities across demographic groups. Critical consideration must be taken to ensure that women are not disproportionately denied loans or offered different terms, incorporating an intersectional perspective. Collaborating with social scientists can help reveal overlooked issues and guide responsible design. More broadly, fintechs must reflect on how institutional priorities shape gender equity efforts. Future research could explore how non-profit institutions approach ML-based lending differently. Regulation also plays a key role.

6.6. Conclusion

By unpacking how fintechs define and measure fairness this chapter shows how deeper social inequalities are often sidestepped – not considered nor prioritized in processes and outcome measurements – and, in many cases, their reproduction through machine learning tools is legitimized. As outlined in Chapter 5, interviewees shared that women are considered better repayers, but are not getting as many loans and receiving loans at lower loan amounts, bringing questions and considerations of fairness to the surface in the interviews. I found that fintechs consider fairness from the perspective of process and outcome. From a process perspective, fairness is pursued through having the machine be unaware of demographics (reflecting Silicon Valley logics to "leave it to the machine") and being available on an app that hypothetically anyone with a smartphone could access. From an outcome perspective, fairness is considered achieved when social impact is achieved and/or models are accurate. In digging into the concept of accuracy, several interviewees revealed that understandings and assessment of model accuracy are linked to pursuit of profit. Across these perspectives of and approaches to fairness, questions of equity and social justice are ignored.

These perspectives and approaches to fairness highlight how fairness in ML is a subjective concept. Meanwhile, a disconnect exists between rhetoric of fairness being achieved and the practice of ensuring or measuring fairness in the model. Fairness tied to accuracy, exposes how both concepts are value-laden, informed by institutional logics, incentives, and priorities. Perceptions of fairness in ML tools are entangled with capitalist priorities aligned with their for-profit creators. Business priorities and broader cultural and institutional logics drive the selection and translation of values into algorithmic design and management illustrating the role of situated priorities. Meritocracy becomes a myth further hidden by market logic: People aren't solely rewarded for merit, and profit-based reasoning obscures and justifies differential outcomes. Do-gooderness provides a moral cover and serves as a blindfold to fintechs themselves, setting the stage for the amplification of gender inequality under pursuit of both profit and scale goals.

However, it does not have to be this way. This chapter reveals that fairness is both a managerial and technical construct. Managers have flexibility in regards to which outcomes are prioritized and what is considered fair, while fairness metrics embedded in the model design and evaluation constrain or shape how those outcomes are operationalized and justified. The two are in constant feedback, each influencing the other. Importantly, as actors begin to recognize how their priorities shape technological outcomes, a reflexive moment emerges that allows a questioning of the neutrality of ML systems and how fairness is defined.

Ultimately, fairness in ML is not neutral or purely technical. It involves value-laden tradeoffs with real consequences. While using profit as the optimization goal isn't inherently wrong in algorithmic lending decisions, it comes with consequences. It can entrench gender inequality when power dynamics and structural inequities in data and institutions are not addressed. All of this occurs without sufficient regulation to guide companies and under opaque models, thereby presenting immense risks. This includes shaping perceptions – and realities – of who is deemed creditworthy, not only by ability and willingness to repay, but by their level of profitability. While a perfect solution may be unrealistic, a better – transparent and inclusive – solution can certainly be strived for.

In Chapter 7 I explore the particular case of a prominent fintech in Kenya to listen to the voices of users themselves, as well as analyze quantitative data around algorithmic-facilitated lending. This chapter delves into the nuances of algorithmic-facilitated lending including their gender impacts and implications.

CHAPTER 7: NEW FINANCE, NEW OPPORTUNITIES – A KENYAN CASE STUDY

7.1. Introduction

This chapter presents a case study on one fintech company with an ML-based alternative lending app in Kenya. It draws on the voices of people using the fintech's app to answer RQ2: *What benefits and challenges do users experience in accessing and using ML-based alternative lending tools, and how do these compare between women and men?* This question informs my overarching research question regarding whether ML-based credit assessment tools reinforce or mitigate gender inequitable access to finance by drawing on insights from users themselves to explore their experiences, by gender. The analysis draws on two datasets, including (a) a dataset of survey data from 342 respondents, and (b) a dataset of 234,740 user reviews retrieved from Google Play. Both datasets reflect self-reported experiences and perceptions of Kenyan users of the same fintech app, rather than administrative records.

I selected Kenya as the case for empirical and strategic reasons: it was the first major market for such technologies and remains one of the most saturated, while gender gaps in digital literacy and mobile internet adoption in the country persist (GSMA, 2024). In addition, the study benefited from a research partnership with the fintech firm, which facilitated survey recruitment in its largest market. Further, Kenya experienced a five-fold increase in people using microfinance and digital credit app providers between 2021 to 2024 (Central Bank of Kenya, 2024), underscoring the importance of understanding gendered experiences and impacts amid rapid growth in digital credit. Without critical attention to gender differences in how people engage with and are impacted by these tools, there is a serious risk women may be left behind in the evolving landscape.

Employing a conceptual framework drawing on gender and development research, feminist economics, and technofeminism, I had several hypotheses. First, women, particularly rural women, experience greater challenges than men (hypothesis 1), which is associated with their relatively lower digital and financial access and literacy (hypothesis 1b). Second, women who access the apps experience greater benefits relative to men, particularly female entrepreneurs (hypothesis 2), associated with their relatively lower access to finance (hypothesis 2b). I employed a multi-method analytical strategy grounded in my conceptual framework, which included descriptive statistics, regression modeling, and mediation analysis to test whether women reported greater challenges or benefits and whether these were shaped by factors such as financial access and digital literacy. I conducted computational text analysis of user reviews and open-ended survey responses to complement and contextualize the survey findings, offering additional insight into user experiences and gendered patterns that may not be fully captured through structured survey data alone.

Key findings reveal a nuanced picture. First, while men and women report similar understandings of the tool and loan process, women more often need help and experience challenges downloading and using the app. These disparities are partially explained by differential access to formal financial services. Secondly, rural women are disproportionately likely to face the challenge of not receiving a loan offer compared to non-rural women and men, a significance which persists when controlling for demographics and not fully explained by formal financial or digital exclusion indicators. Small loan size is also significantly associated with not receiving a loan offer but does not mediate the effect of being a rural woman.

Third, while men and women experience positive impacts from using the app and accessing loans, gender differences emerge. Female entrepreneurs in particular exhibit higher benefits than other women and men in increased hopefulness, greater sense of financial control, and enhanced financial decision-making. Among female entrepreneurs, impacts are partially mediated by a lack of other loan options, suggesting that women may perceive greater value due to fewer available alternatives.

Fourth, exploratory findings provide an important nuance: women in the sample arrive to the app more financially excluded than men in the sample and receive smaller loans via the app, while reporting better repayment behaviors. The app does indeed provide access to finance to people who are otherwise outside of formal financial folds. Still, both women and men who come to the app are better off than the average Kenyan population in regards to financial inclusion.

Taken together, these findings suggest that while ML-based credit assessment apps offer real benefits to women and men, gender differences persist. Without deliberate approaches to understand and address gender disparities, these financial technologies risk perpetuating existing inequalities in financial access. These insights underscore the need for fintechs to adopt a gender lens in both app design and algorithmic decision-making.

The chapter is structured as follows. I begin with my conceptual framework and hypotheses. I provide an outline of the context of Kenya, including financial and digital inclusion trends. I then discuss my data, variables, and analytical approach. Results are presented in two sections. First, I begin with descriptive statistics of my survey data to explore gendered patterns in benefits and challenges. Then, I turn to regression analysis to test key hypotheses, followed by mediation models examining the mechanisms driving these patterns in the survey data. I weave in results from the text analysis conducted on the user review data and open-ended survey responses. I provide a discussion on findings before concluding the chapter.

7.2. Background and Context: The Case of Kenya

Kenya is recognized as a pioneering country in mobile money and financial services, and continues to lead in digital credit innovation, including ML-based credit assessment. Indeed, Kenya was one of the first markets for ML-based credit assessment technologies in LMICs, which have since taken off in the country. This growth is linked to several trends in the last twenty years: the growth of mobile money, increasing proliferation of smartphones, and the spread of formal credit scoring. These trends have impacted financial inclusion, while creating new market opportunities including as it relates to fintechs offering alternative lending apps. However, these advances take place alongside persistent gender gaps in digital and financial inclusion. This section provides an overview of Kenya's digital and financial inclusion landscape, with a focus on gender disparities, borrowing behavior, and the emergence of ML-based credit scoring systems. Understanding these dynamics is important to interpret the findings of my study and situate Kenya as an ideal case for my research.

7.2.1 Financial inclusion in Kenya

In Kenya, 83% of men have a formal financial account, while 75% of women have formal accounts (World Bank, 2021). Formal financial access increased by 3% between 2021 and 2024, which was largely driven by digital technology (Central Bank of Kenya, 2024). The 8 percentage point gender gap in formal account ownership is less than the broader sub-Saharan Africa (SSA) region, where there remains a 12 percentage point gender gap in account ownership (twice the developing economy average). Outside of gender, other key factors correlated to not owning a formal account in Kenya include lower education and living in rural locations (Central Bank of Kenya, 2024).

In Kenya, 73% of women and 80% of men have borrowed money. Informal borrowing from friends and family is the most common source of credit for both men and women, while borrowing from community savings groups remains an important source of credit, particularly for women (World Bank, 2024). Mobile money accounts play important roles in borrowing, with 37% of women having borrowed from a formal institution or mobile money account versus 42% of men (World Bank, 2024).

Mobile money has played an important role in enhancing financial access and inclusion in Kenya, where it has been popular since the early 2010s and a robust mobile money market exists. The popularity of mobile money is linked to the history of M-Pesa in the country, which is a mobile phone-based service for sending and storing money offered by Safaricom, Kenya's largest mobile service provider (Morawczynski & Pickens, 2009). Mobile money has brought more women into financial systems than banks have and contributes to the lower gender gaps in account ownership in Kenya. Mobile money account ownership among women is at 66%, whereas for men it is 71% (World Bank, 2021). Meanwhile, having an account at a financial institution account in 2021 was lower at 45% for women (down from 52% in 2014) and 57% for men (down from 59% in 2014) (World Bank, 2021). This trend reflects how phones have provided fertile spaces for access to financial services, with mobile money becoming a more popular way for people to save.

There has been an increase in mobile money, with 52.6% of Kenyans using mobile money daily as of 2024, more than doubling from 2021 (Central Bank of Kenya, 2024). This is an ongoing trend that was accelerated following the COVID-19 pandemic, which was linked to an acceleration of digital payments and mobile money usage (World Bank, 2021). Mobile money supports access to financial services including being able to make and receive payments (e.g., peer transfer, wage deposits), save money, and access financial borrowing opportunities (World Bank, 2021). Many mobile money providers provide a digital loan product, with digital credit being the most prevalent non-payment service offered, with the majority being offered after 2019 (CEGA, 2024). As of 2021, 30% of Kenyans used mobile money to borrow (World Bank, 2021). Taken together, this information illustrates the growing role that mobile money and digital financial services play in lending and accessing financial services.

Still, there remain many unbanked people. Reasons unbanked people lack financial accounts can include not having enough money to justify having an account, lack of documentation to open a bank account, distance to physical banks, and expenses of financial services. Women are more likely than men to cite lacking documentation as a barrier, as well as not having an account because another family member has one (World Bank, 2024). Technology has provided a way to address some of the barriers to formal financial access, while introducing new ones for digital financial services. For example, not having a phone presents a persistent challenge in accessing digital financial services.

7.2.2. Gender gaps and challenges in digital financial inclusion

Fintech solutions can suffer from persistent global gender differences in digital inclusion and literacy, as well as financial literacy (Hasler & Lusardi, 2017). Surveys and research by GSMA reveal that the gender gap in mobile internet adoption is closing, yet large gaps persist in sub-Saharan Africa, with further amplification in rural areas (GSMA, 2024). In Kenya, while both men and women own mobile phones (92% and 91%, respectively), the gender gap in mobile Internet use is large: only 39% of women use mobile Internet versus 58% of men (GSMA, 2024). Meanwhile, although 42.1% of the population is considered to have high financial literacy, there is a gender gap: 45.8% of men reported being highly literate compared to 38.3% of women (Central Bank of Kenya, 2024). The combination of these different gender gaps make it a particularly interesting case to study for my research.

Challenges persist in accessing and using mobile money accounts and digital financial services. In particular, using digital financial services requires digital literacy skills to navigate user interfaces, manage passwords and accounts, and more. Other challenges include, for example, lack of transparency from providers about fees and other terms of services and mobile app fraud (Chalwe-Mulenga & Duflos, 2021). Nearly one-third of mobile money account owners in SSA cannot use their accounts without help, with many more women requiring help than men (World Bank, 2021). Still, from a broader market perspective, sub-Saharan Africa is one of the largest digital credit markets (with Kenya as a leader), according to research by the UC Berkeley Center for Effective Global Action, with approximately $400 billion in credit offered in 2019 (CEGA, 2024).

7.2.3. Credit scoring and the growth of digital lending apps in Kenya

While credit assessment and credit bureaus emerged in the US and later England, they started to take hold in LMICs in the early 2000s. The first credit bureau in Kenya was established and licensed by the Central Bank of Kenya in 2010, called the Credit Reference Bureau Africa (now TransUnion Africa). In 2013, regulations were implemented allowing credit information to be shared by banks and microfinance institutions with credit bureaus (Mungiria & Ondabu, 2019). Prior, credit data sharing was done through the Central Bank of Kenya with banks providing data reports. Discussing the importance of credit bureaus, Kamau Kunyiha, CEO of Creditinfo Kenya, says: “One of the obstacles for credit access is a lack of information about borrowers… Credit bureaus services are crucial – they provide the data financial institutions need to provide individuals and businesses access to essential financial products and encourage a healthier financial ecosystem" (TransUnion Africa, 2022). In Kenya (as well as other LMICs), credit bureaus were deemed not necessary due to technological limitations prior to the 2000s. TransUnion – along with other prominent US and UK credit bureau entities – played an important role to inform the advance of credit bureaus in the countries, including Kenya.

The growth of digital lending apps in Kenya began in 2012 with M-Shwari. M-Shwari, a banking service from Safaricom that offered micro-savings and lending capabilities to users of M-Pesa, helped spur the growth of digital lending apps in Kenya. From 2012, various players have offered mobile loans via apps that use proprietary credit scoring algorithms. Outside of M-Shwari, other prominent players include, for example: KCB M-Pesa, Branch, Tala, and Equitel. It is not only fintechs that are using ML for credit scoring. In early 2025, TransUnion and FICO announced a partnership for a new proprietary risk scoring algorithm analyzing "over 145 data sources", illustrating how the ML-based credit scoring is expanding beyond fintech company solutions (FICO, 2025).

While there are limited statistics on how large the market is for ML-based alternative lending and credit scoring apps in Kenya, there are helpful indicators of the size and the popularity. The Central Bank of Kenya 2024 survey finds that 8.8% of respondents use microfinance institutions and/or digital credit app providers, which totals over 5 million people (Central Bank of Kenya, 2024) and is a fivefold increase from 2021. These statistics do not include digital loans provided on mobile money apps or via neobanks, which are digital banks that operate solely online without physical infrastructure. Neobanks are quite new in Kenya, with the first neobank in the country in 2023 launched by a fintech that is a leader in the ML-based credit assessment space: Branch International (Ukibe, 2024). Overall, in Kenya, 64% of the population is taking out credit of some kind, up from 60.8% in 2021, with the increase attributed to the rise of app-based digital loans and mobile money credit. The apps take advantage of high mobile money penetration alongside a growing number of digital credit providers (Central Bank of Kenya, 2024). The popularity of these apps is also reflected in the number of downloads and reviews for leading providers. For example, Tala has over 10 million downloads and nearly 470,000 reviews for its Kenyan app alone.[17] Branch has over 50 million downloads and 2.33 million reviews across its markets of Kenya, India, Nigeria, and Tanzania.[18]

Financial health indicators in Kenya illustrate why demand for these types of apps providing micro and small loans may be popular. Financial health is determined to be low, with only 18.3% of Kenyans considered financially healthy in 2024. There has been a recent decline of people able to invest in or save for the future, with only 68% of Kenyans reporting saving money (down 6% from 2020) as households prioritize meeting day-to-day needs (Central Bank of Kenya, 2024). From a gender perspective, women are more likely to rank school fees as their biggest financial worry, versus men ranking medical expenses the highest (World Bank, 2021). Digital lending tools can help address financial needs and support financial well-being. Importantly, digital lenders in developing economies are significantly less regulated than fintech lenders in developed countries, while the loans tend to be smaller with shorter duration and high interest rates (A. Y. Chen et al., 2025).

7.2.4. Study context: Motivation for Kenya case selection and the fintech partner

Taken together, Kenya is the ideal country as a case study for this research. Its history as a leader in mobile money and ML-based credit assessment, persistent gender disparities in digital and financial inclusion, and the ongoing growth and demand for digital loan services in the face of low financial health make it strategically relevant for my research. Importantly, my research partner's largest market is in Kenya, which made it particularly of interest to my research partner and enabled the practical implementation of the study.

[17] Retrieved on March 9, 2025:
https://play.google.com/store/apps/details?id=com.inventureaccess.safarirahisi&hl=en_US
[18] Retrieved on March 9, 2025:
https://play.google.com/store/apps/details?id=com.branch_international.branch.branch_demo_android&hl=en&gl

The fintech partner ("the Lender") administers digital loans using an app with ML-based credit assessment. Applicants apply for loans through the Lender's mobile app, which they download from Google Play. After downloading and opening the app, users are able to complete a simple loan application, after which they grant the Lender access to mobile phone data. Borrowers don't request a certain amount; rather, the Lender determines the credit the borrower is eligible to receive using ML algorithms, as well as repayment terms.[19] The Lender has administered more than 3 million loans in Kenya and operates in several other countries globally. The Lender has built a reputation as a reliable lender in Kenya and is purpose-driven with a focus on enhancing access to finance to people. The extensive user base in Kenya provided a valuable foundation for exploring my research question. The Lender remains anonymous per our research agreement and therefore I provide no further information about the Lender.

7.3. Conceptual framework and research hypotheses

7.3.1. Conceptual Framework

This chapter draws on gender and development research to recognize the structural differences that impact women's access to and use of technology. It employs a feminist economics lens to explore how users experience impacts of the tools, with attention paid to how relative impacts may be greater for women (particularly among those more credit constrained), while gendered tensions can persist in the tools' access and use linked to gendered structural divides in digital and financial access and literacy.

My conceptual framework emphasizes that women, particularly those who are rural, face greater challenges in access and use of ML-based credit assessment apps. These challenges stem from two areas. First, women experience persistent structural inequalities, including women's relatively lower digital access and literacy (GSMA, 2025), as well as lower financial literacy (Hasler & Lusardi, 2017). Further,

[19] No further information is provided about the Lender as I have signed an NDA and they are to remain anonymous per our research agreement.

scholars point to women's "triple role": the productive, reproductive and community roles (Moser, 2012). The ILO finds that women's disproportionate share of caregiving responsibilities serve as the main barrier globally to women's entry and participation in the labor force, while greater caretaking responsibilities are also linked to more part-time work and informal employment (ILO, 2024). Women are not a monolith, with various other aspects of identity – such as class, ethnicity, location, and age – interacting to inform people's lives over time and place (Crenshaw et al., 1996; Parpart et al., 2000). Rural women, for example, face multiple layers of disadvantage in the labor market and are often present in the informal economy, while women with more caregiving responsibilities tend to have lower education (ILO, 2024). Women's intersecting identities also inform their access to and experiences of technologies. This is evidenced by women who are from low-income backgrounds and rural residents experiencing greater exclusion from formal financial systems (Kim & and Duvendack, 2024). Further, rural women are least likely to be digitally included (GSMA, 2025).

Secondly, gendered differences in challenges can stem from the design of the technologies themselves, including the app and the ML technology that remains "blind" to gender and privileges a default user that is male. This mirrors Wajcman's assertion that technologies are not neutral but shape and are shaped by gender (Wajcman, 2006). Similarly, while technology can and does bring important benefits, technologies often fail to deliver on their promises due to not recognizing the role of power and hierarchies – including related to gender – that impact how benefits are distributed (Noble & Stalder, 1998). Being "blind" to gender is not new in technology design and is also found in development interventions, linked to a lack of awareness on women's time spent on unpaid and paid work, as well as how unequal gender relations structures economic processes (Elson, 1995). ICTs can both create new opportunities to circumvent traditional restrictions but still remain embedded with norms and power relations in the culture (Philip, 2018). Although my statistical models focus on observable differences in challenges and benefits, in the discussion of this chapter I explore how gendered experiences of the apps are linked to technological design and management choices by drawing on Social Shaping of Technology (MacKenzie & Wajcman, 1999) and technofeminism (Wajcman, 2006).

My conceptual framework also explores how women may experience perceived greater benefits relative to men. Feminist economists highlight how access to and control over resources – such as property, land, and credit – are tied to economic agency and women's empowerment (Kabeer, 1999). Empowerment is the ability to make and act upon strategic life decisions that were previously denied and includes three key components: resources, agency (ability to make and act on important choices), and achievements (outcomes) (Kabeer, 1999). Women's empowerment and confidence is linked to their resources and ability to make social and economic decisions (Horton, 2017). Even limited access to resources can lead to empowerment outcomes, particularly for marginalized women (Agarwal, 1995). Women face lower access to formal financial services globally (World Bank, 2021), with female entrepreneurs facing persistent credit discrimination (de Andrés et al., 2021). Women may therefore experience heightened perceived benefits by gaining access to financial resources when historically excluded from formal financial systems, which is particularly the case for women entrepreneurs.

7.3.2. Existing Research

Existing research on the benefits of ML-based credit assessment tools tends to focus on welfare impacts of access to loans facilitated through the tools and is mixed, with limited gender analysis. Research in Nigeria finds that those who borrowed via an ML-based alternative lending app had a modest increase in financial health and modestly positive effects on subjective well-being, but other measures of welfare were insignificant (Björkegren & Grissen, 2020). A study using data from another leading fintech providing ML-based alternative lending in Kenya found that access to digital credit improved borrowers' financial well-being, with greater impacts among borrowers with limited access to credit and those who take loans for business purposes (A. Y. Chen et al., 2025). Other studies show that loans via ML-based alternative lending help cope with unexpected income shocks (Brailovskaya et al., 2021; Suri et al., 2021), but loans may be too small for business growth (J. Robinson et al., 2023). Other researchers illustrate unintended consequences, including high interest rates and debt traps resulting from ML-based alternative lending tools (Qureshi, 2020; J. Robinson et al., 2023). In Kenya, there have been several regulations to curb digital lending, particularly linked to high interest rates and predatory lending practices including from bad actors (Mutua, 2021). Across these studies that assess impacts for those who do access loans, impacts are not found to be different among men and women, or gender differences are not explored. Intersectionality analysis remains lacking.

At a higher level, there is a fintech gender gap impacting who is able to access and use the apps to begin with. In a global study of 28 countries, 21% of women use fintech products compared to 29% of men (S. Chen et al., 2023). This can be linked to differences in digital inclusion and literacy, as well as other structural constraints (Johnen & Mußhoff, 2023). Furthermore, fintechs tend to operate under "gender neutral" approaches: The IFC (2024) conducted a survey of fintech firms from 17 countries to find that 68% of fintechs do not tailor design and delivery services for women. Perhaps relatedly, women constitute less than 25% of fintechs' customers. At the same time, executives in the majority of fintech firms consider women to be more loyal, less risky, and have better repayment rates compared to men (International Finance Corporation, 2024). In short, fintech products are not being designed with a gender lens and are not reaching as many female consumers, despite the economic potential of their inclusion.

Relatedly, research conducted on digital credit in Kenya found that formal digital credit led to an increase in the gender gap in financial inclusion. Gender differences were largely attributed to gender differences in socio-economic variables combined with "a lack of contract term heterogeneity in that market" (Johnen & Mußhoff, 2023). This implies the digital credit market did not adapt to gendered differences illustrating the importance of supplier-side attention to gender differences in use and access.

7.3.3. Research Hypotheses

Drawing on my conceptual framework, I have the following hypotheses.

- Hypothesis 1 (H1): Women experience more challenges than men in accessing and using ML-based credit assessment apps. These challenges are more pronounced among rural women.
- Hypotheses 1b (H1b): These greater challenges are associated with lower levels of digital and financial inclusion and literacy among women.

H1 and H1b are grounded in the understanding that women, particularly those in rural areas, face persistent structural barriers to digital and financial inclusion and literacy. Drawing on feminist economics and gender and development literature, I expect these structural constraints, alongside gender norms, to inform women's experiences with ML-based credit assessment apps. Specifically, women – particularly in rural locations where structural constraints tend to be larger – may encounter greater challenges in accessing and using these tools, while the design and deployment of ML-based technologies may implicitly privilege male users. I therefore expect that women, especially rural women, will report more difficulties engaging with these tools, and that these difficulties will be linked to lower levels of digital and financial inclusion.

- Hypothesis 2 (H2): Among those who gain access, women, particularly those self-employed, experience greater relative benefits from ML-based credit assessment apps compared to others.
- Hypothesis 2b (H2b): Greater benefits are associated with women's relatively more limited access to financial services.

H2 and H2b reflect the expectation that women who access and use ML-based credit assessment apps, especially those who are self-employed, may report relatively greater perceived benefits once they do gain access. Drawing on feminist economics, incremental access to resources such as financial tools can be empowering for women who have historically been excluded from formal systems. This could be particularly true for self-employed women, whose credit needs may be unmet by traditional banks or other informal sources. As such, I expect that women, especially female entrepreneurs, will perceive stronger gains from using the tools. These benefits may be amplified by their more constrained baseline conditions: for women with fewer financial alternatives, gaining access to even modest credit via ML-based apps may feel especially impactful. Patterns of financial exclusion are not only relevant to understanding access but may also help explain why some women report greater relative gains.

While H1 and H2 focus on associations between gender and reported challenges or benefits, H1b and H2b propose potential mechanisms; namely, lower levels of financial and digital inclusion as contributing factors. Directional expectations are grounded in theory. However, all analyses are based on observational data and should be interpreted with caution, as they do not establish causal relationships.

While the hypotheses focus on gendered patterns in access, use, and outcomes, the study also explores broader user-reported benefits and challenges to provide a more comprehensive understanding of how ML-based lending tools are experienced. These insights complement the hypothesis-driven analysis.

7.4. Data and Variables

7.4.1. Data Sources

7.4.1.1. End user survey

I conducted a survey with 342 respondents who are users of the Lender's app in Kenya. In order to distribute the survey to end users of the app, I developed a partnership with the Lender, whose reviews I also scraped from Google Play.[20] After several months of discussion and negotiation, as well as signing a Non-Disclosure Agreement (NDA) with the Lender following multiple back-and-forth emails with the legal team, the partnership was solidified in fall 2023. In April 2024, the survey was implemented (see Survey Questionnaire in Appendix 2). I met with a representative of the Lender over a period of several months prior to the survey implementation to discuss and refine the survey approach to meet my research goals.

[20] The partner was aware of the review data I scraped and utilized. We also discussed whether they could provide me with gender-disaggregated review data, but this is not something they collect.

There were several steps taken with the Lender prior to the launch of the survey. First, I shared the survey with the Lender so they understood the types of questions being asked and the purpose of the survey. I developed a recruitment plan and approach that outlined the goal to reach 300 respondents with the following user profiles: (a) *Gender:* At least 50% women (# = 150 women); *Location:* At least 30% rural (# = 90 rural; # = 210 urban); and *First-time users:* At least 30% first time loan recipients of the app (# = 90 first time; # = 210 repeat). All respondents should have used or accessed the app in the last three months. I provided text for the push notification that was reviewed by the fintech partner. We piloted the survey with several Kenyan staff of the Lender, as well as several customers of the Lender, who were compensated for their time. The feedback was then incorporated into a final version of the survey on Qualtrics. Part of the feedback included having a version that was in Kiswahili. Therefore, I hired a translator to translate the survey into Kiswahili and a translated version was made available on Qualtrics for all respondents to select if desired. Overall, all respondents used the English version of the survey.

For participant recruitment, the invitation to participate in the survey was shared by the Kenyan office of the Lender, who shared the invitation with end users who met the inclusion criteria through a push notification sent to their mobile phone. The Lender estimated a 2% response rate, based on prior surveys, and therefore we planned to message approximately 15,000 people. The Lender sent out three rounds of push notifications between April 16th and April 29th, 2024. By sending out the push notifications in waves, I sought to gradually reach my sample of 300 respondents. This was recommended by the Lender, as they anticipated a potentially higher response rate due to the survey being paid. Given my demographic respondent goals (e.g., 50% female borrowers), the Lender sent out push notifications at random while also reflecting the percentages I sought, as opposed to sending out push notifications to app users fully at random. This was important as the Lender has fewer female borrowers relative to men, low levels of rural borrowers relative to urban and peri-urban, and fewer first-time borrowers relative to repeat borrowers. I also added limits in Qualtrics so that the survey would not continue for folks once different demographic goals had been reached. In the second round of push notifications, the Lender noted that the sample size for rural versus urban was smaller in the last three months and expected I may not reach the desired sample size. We therefore increased the inclusion criteria to users that had accessed the app in the prior four months. Within the third round of push notifications I reached my sampling goal. Within this round, I reached 470 responses (indicating an approximately 3.1% response rate); however, given the limits in Qualtrics, 98 of those responses were stopped when the potential participant entered in their demographic information and Qualtrics alerted the individual that they could not continue. An additional 30 responses were removed due to incomplete information or repetitive entries.

End users who wanted to participate clicked on a link from the push notification to go to a Qualtrics form with the inclusion criteria and consent form. The first page of the survey in Qualtrics was on informed consent. The online consent form included: general information on the study and survey and remuneration provided, affirming they have a choice to take part, how data will be used, who will have access to the data, who has reviewed the research, procedure for complaints or concerns, my contact information. In regards to access to the data, I made clear that the Lender would not have access to the data. Rather, only myself and my supervisors at University of Oxford would have access to the research data. The form required participants to certify that they are over 18 years old. For the survey, there was remuneration for participants' time in completing the survey equivalent to 3 US dollars (USD 3), which was identified based on the time required to complete the survey and recommendation remuneration rate of other Oxford researchers conducting surveys in Kenya. When a survey was completed, participants were prompted to submit their name and phone number to receive payment. The payment was made via a mobile payment service (MPesa) using a digital money transfer service (Sendwave). I maintained participants' anonymity and confidentiality by destroying names and phone numbers after payment was processed.

Overall, 342 complete surveys were collected for my sample, with 47% of respondents being women. An overview of select demographics of the survey respondents are outlined in table 6. In terms of location, 13.7% of respondents self-identified as rural, 36% as living in a "town", and 50% as living in a large city. Meanwhile, 11% were first-time customers. In addition to the information outlined in Table 6, I collected information on the loan purpose and financial inclusion status, which I further discuss in the findings section. In Table 6, I include information for the overall Kenyan population to see how the survey sample is similar (or not) to broader demographics in the country. Notably, more than 50% of respondents, both male and female, reported having a university degree, which is much higher than the overall Kenyan population. This likely reflects two dynamics: first, the characteristics of users who access and engage with the Lender's app itself, which requires digital literacy and smartphone access; and second, potential response bias (further discussed in limitations in 7.8). While it's possible that more educated users were more likely to respond, the demographic skew likely reflects the broader user base of the app more than selective response alone. Per my agreement with the Lender, I presented results of the survey to an internal team of the company in November 2024. Anecdotally, when I presented the demographics of the respondents to the Lender they noted that the demographics reflected their understanding of their user base (except the more equal percentages of men and women in the sample as aforementioned), affirming that the skews likely reflect the broader user base more than selective response alone.

Table 6. Survey respondent demographics

	Male (%)	**Female (%)**	**Total (%)**	**Overall Kenyan pop (male, %)**	**Overall Kenyan pop (female, %)**
Married	65.93	36.88	52.03	45.8[A]	55.4[A]
Single / widowed / divorced	34.07	63.12	47.97	N/A	N/A
University degree	52.75	54.37	53.2	3.5[B]	3.5[B]
Completed secondary education	15.38	15	15.2	N/A	N/A
Some secondary or less	8.75	8.8	8.78	N/A	N/A
Vocational or technical training	23.08	21.88	22.5	N/A	N/A
Youth (18-34 yrs old)	56.6	70	62.4	34.7[C]	34.7[C]

Aged 35+	43.4	30	37.6	N/A	N/A
Rural	13.81	13.84	13.82	70[D]	70[D]
Urban and peri-urban	86.19	86.16	86.18	N/A	N/A
Employed	46.70	33.75	40.64	42.6[E]	25.6[E]
Self-employed / running own business	32.46	33.12	31.87	57.4[F]	74.4[F]
Casual worker or agriculture	13.18	21.87	17.25	N/A	N/A
No income	5.85	4.95	6.88	N/A	N/A

[A] Ages 15-49, from (*Marriage Rates by County - Prevalence of Monogamous and Polygamous Marriage*, 2022)
[B] From the 2019 Kenya consensus reported (not gender disaggregated) by (Statista, 2019)
[C] Estimated and provided (not gender disaggregated) in (National Council For Population and Development, 2017)
[D] Rural population overall, not by gender, from (World Bank Group, 2023)
[D] Wage and salaried workers, from (World Bank Group, 2023)
[F] Percent of total employment, from (World Bank Group, 2023)

In building out the survey, I used several approaches and resources. First, I had a question for respondents to select different financial activities to measure different aspects of financial inclusion, the categories for which were drawn from the financial inclusion definition of the World Bank (World Bank, n.d.). My questions around sources of income and purposes for taking out one's latest loan were drawn from surveys on digital credit by CGAP (CGAP, n.d.). The impact categories focused on women's empowerment and financial empowerment indicators, which were informed by J-PAL and CGAP (Deshpande & Koning, 2023; J-PAL, n.d.). Categories for challenges were identified through considering challenges throughout the lifecycle of accessing and using the app (e.g., opening the app and usability, receiving loan offers and monies, repayment, and customer service).

7.4.1.2. User reviews retrieved via Google Play

In addition to the survey data, I gathered review data by scraping user reviews from the Lender's page on the Google Play store. The particular fintech (i.e. my research partner) had over 1.8 million reviews as of August 2024. Android devices dominate the market in Africa and Asia (*Mobile Operating System Market*

Share Africa, 2024; *Mobile Operating System Market Share Asia*, 2024) and Google Play is where these apps are located for download, hence I am using Google Play.

I scraped all reviews that were available in January 2025. The dataset includes reviews provided on a date ranging from April 2020 to August 2024. In total, 234,740 reviews were scraped; of which, 38,344 had names associated with the review and I was able to assign gender in 18,799 reviews. Details of the computational methods, the various Python packages used, and the accompanying reasoning confirming my approach as widely used and appropriate is outlined in my analytical approach under section 7.5.2.

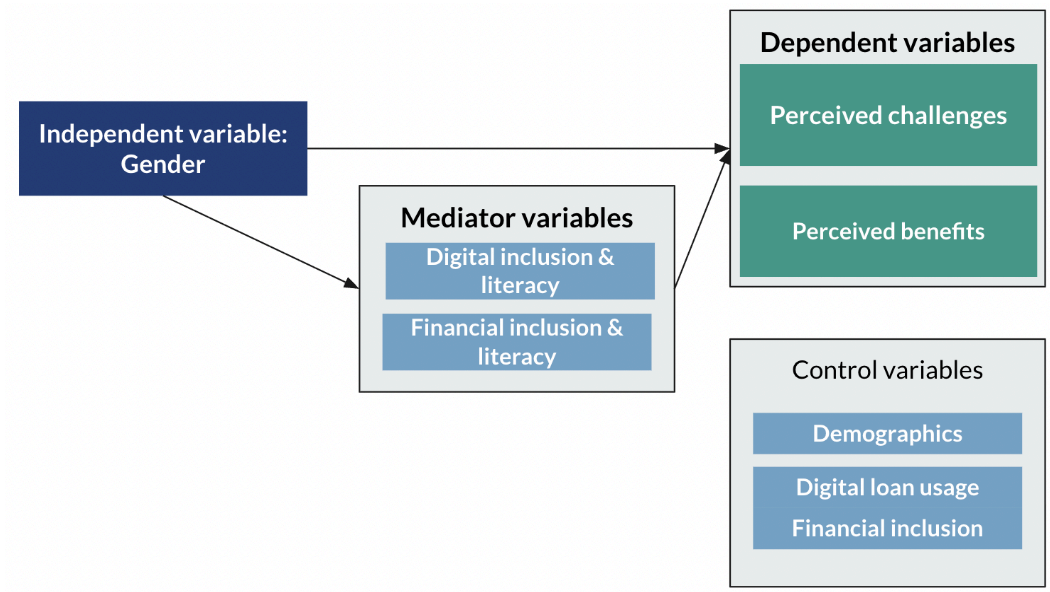
7.4.2. Independent, dependent, and control variables

7.4.2.1. Overview of variables

In this chapter, the main independent variable is gender, coded as a binary variable. Although nonbinary was included as an option in the survey, no respondents identified as such. I therefore coded gender as binary; however, I recognize that gender is a spectrum. The dependent variables relate to (a) perceived challenges in accessing and using the app, and (b) perceived benefits from app use and credit access. Benefits are measured through Likert-scale items assessing both financial benefits (e.g., one's ability to meet basic needs) and personal empowerment benefits (e.g., improvements to self-confidence, financial well-being, and financial decision-making ability). Challenges are measured through a combination of binary items regarding access to loans, using the app and repayment issues (e.g., whether the user received a loan offer, needed help downloading or using the app, didn't receive payment reminders) and Likert-scale items (e.g., the extent to which one understood the loan approval process, understood all terms and conditions, and had no trouble making payments).

I also included several control variables in the following categories: (a) demographics, (b) digital loan app usage, and (c) financial inclusion status. Demographic variables included: employment status, rural versus urban location, education level (university degree versus other), age group (youth versus non-youth), and marriage status. Variables related to digital loan app usage included whether this was the first loan taken via the app and if the user has multiple digital loan apps. Variables for financial inclusion included whether users have accessed any formal financial services and whether they have other loan options. These digital loan app usage and financial inclusion indicators are included not only as control variables in regression models, but as mediators in the mediation analysis designed to test underlying mechanisms for gender differences in challenges and benefits. The relationship between the variables and links to my hypotheses are captured in Figure 2.

Figure 2. Visual of conceptual model and variables



7.4.2.2. Variable transformation, composite construction, and interaction terms

To reduce dimensionality and avoid multicollinearity in regression analysis, I re-coded multi-category demographic variables into binary indicators. For example, while respondents could select from several income sources, I created a binary variable for employment status (1 = formal employment, 0 = all other income types). Similarly, rural location is coded as a binary variable (1 = rural, 0 = urban or peri-urban), education is represented by a dummy for holding a university degree (1 = university degree or higher, 0 = less than university). This approach allowed me to retain theoretically meaningful contrasts – such as employment status, rural location, and educational attainment – while maintaining model parsimony and statistical robustness (i.e. keeping models simple and interpretable, while maintaining explanatory power). These simplified indicators also reflect the kinds of socioeconomic distinctions that emerge as important considerations or proxies in ML-based credit scoring models, as identified in both the literature and interview data.

I also constructed composite variables for key mediators and control variables aligned with my conceptual framework and analytical strategy. In particular, *formal financial inclusion* is a binary variable coded as 1 if a respondent reported any of the following: receiving a loan from a formal financial institution, opening a savings account with a formal institution, or holding formal insurance. This measure captures engagement with the formal financial system. Similarly, the variable *loan purpose work* is coded as 1 if the respondent reported that their most recent loan was used for a productive activity (e.g., starting a business or side hustle, buying inventory or stock). These composite indicators allow me to capture key representations of complex constructs that are theoretically and/or practically meaningful. These composite indicators reflect constructs prioritized either in my hypotheses (e.g., financial inclusion) or identified as relevant in the literature (e.g., work-purpose loans).

Finally, in line with my hypotheses and to explore how gender subgroups may experience the outcomes, I also created interaction terms to use in the analysis, including self-employed women and rural women. These interactions allow me to explore whether relationships between gender and outcomes (i.e. benefits, challenges) vary across subgroups. See Appendix 3 for an overview of all variables, including re-coded multi-category variables to binary variables, composite indicators, and interaction terms.

7.5. Empirical strategy and analytical approach

The analysis followed a multi-method, staged approach integrating descriptive statistics, regression modeling, and mediation analysis. Computational text analysis complemented my investigation of gendered experiences of ML-based lending in Kenya. The analysis was grounded in the theoretical expectation and my hypotheses that women face more pronounced challenges in accessing and benefiting from ML-based credit assessment apps, while also potentially experiencing greater benefits. These expectations shaped the analytical design: interaction terms (e.g., female × rural) and mediating pathways (e.g., digital or financial inclusion) were selected based on my conceptual framework, not data-driven exploration.

After descriptive statistics, regression analysis tested systematic associations with gender. First, I tested whether gender was significantly associated with reported outcomes (Model 1). Where a significant gender gap was observed, I introduced demographic, digital, and financial controls (Models 2 and 3) to assess whether the association persisted or was potentially mediated by structural factors. This allowed me to identify whether disparities were explained by differential access to enabling resources, rather than gender alone. Building on this, I conducted formal mediation analysis to test whether specific mediators – identified in my hypotheses – accounted for observed gender differences. This is my main analytical strategy, further detailed below.

I also employed computational social science methods, specifically text analysis, on the large user review dataset, which offered unstructured data for complementary insights related to my hypotheses. This text analysis combined both exploratory and comparative components: I identified topics that emerged organically from the unstructured user reviews, and also examined whether and how topics that surfaced in the survey findings were reflected in the review data. This approach enabled me to generate new insights from the reviews while also assessing the extent to which patterns identified in the survey were echoed in unsolicited user accounts. I also conducted text analysis on open-ended survey responses to contextualize and deepen interpretation of the structured survey results. Finally, alongside the hypothesis-driven analyses, I include a set of exploratory findings from the survey data that provide context on app users and their loan behaviors.

7.5.1. Main analytical strategy

The analysis was conducted in Python using a combination of open-source libraries: *pandas* was used for data cleaning, organization, and manipulation; *NumPy* supported numerical operations and array handling; and *statsmodels* was used for statistical modeling and diagnostics, including Variance Inflation Factor (VIF) calculations for assessing multicollinearity, regression analysis, and mediation analysis. In addition, scipy.stats was used to conduct non-parametric and inferential statistical tests (e.g., chi-square tests, Mann-Whitney U tests), to compare distributions and associations between categorical and ordinal variables. All analyses were conducted in a local Python environment using Jupyter notebooks, with result tables exported for visualization. I took the following steps:

- *Preprocessing*: I began with a series of data preprocessing steps. Initial review of survey responses was carried out in Excel to identify and remove incomplete submissions. Following this, I imported the cleaned dataset into Python for further processing. In this phase I created the interaction terms to test intersectional hypotheses, composite variables, and re-coded multi-category variables to binary variables.
- *Multicollinearity testing*: I used VIF tests to assess multicollinearity among independent variables, including key interaction terms (i.e. rural female, self-employed female) and ensure the robustness of the regression models. I found low multicollinearity across the coefficients illustrating that predictors are sufficiently independent from each other to allow meaningful interpretation. All variables in the model had VIF scores below 2.2, with the highest being 2.14 for the female rural interaction term (see Appendix 4, Table 1). These results suggest that multicollinearity is not a concern and that the estimates from the regression models are likely to be stable and interpretable. I also created a correlation matrix (see Appendix 4, Table 2), which shows pairwise Pearson correlation coefficients among the independent variables. While most correlations fall in a low to moderate range (generally below +/- 0.5), one correlation exceeds this threshold: 0.66 between rural location and the interaction term rural female (i.e. female × rural),

which is expected given the interaction term's construction. The matrix therefore supports the VIF results in demonstrating low multicollinearity.

- *Descriptive statistics*: After my initial steps, my first stage involved descriptive statistics, examining patterns in reported benefits and challenges, disaggregated by gender. I summarized outcomes using means, standard deviations, and tests of significance, and highlighted gendered differences in challenges and benefits where statistically significant.
- *Logistic regressions*:
    - I used ordered logistic regression for ordinal scales (e.g., Likert items) and binary logistic regression for binary outcomes. Rather than aggregating items into composite indices (e.g., overall benefits, overall challenges), I analyzed each survey item individually. This is because the benefit and challenge items measure different kinds of experiences. For example, technical difficulties with app use are distinct from customer service issues and reflect different types of issues. Creating indices would have obscured these meaningful distinctions.
    - For each item, I followed a sequential, theory-guided model approach by incrementally adding variables in stages only when a significant relationship was observed at baseline. This includes: (1) a baseline model including only gender. If model 1 showed a statistically significant effect, I proceeded to: (2) a model adding demographic controls, and (3) a full model that additionally included digital app usage and financial access controls. This approach reflects principles of model parsimony and helps minimize overfitting (Burnham & Anderson, 2002), while aligning with theory-driven model building strategies (Hosmer et al., 2013).

The general model structure is:

$$Y_i = \alpha_0 + \alpha_1(\text{Female}_i) + \alpha_2(\text{Demographics}_i) + \alpha_3(\text{Digital/Financial Access}_i) + \varepsilon_i$$

where Y_i is the reported challenge or benefit (measured as an ordinal or binary variable), $Female_i$ is a binary indicator for gender, $Demographics_i$ includes demographic variables, and $Digital/Financial\ Access_i$ includes my financial and digital inclusion variables.

While these models control for a range of theoretically relevant variables, the analysis is based on cross-sectional data. The results should be interpreted as correlations, not causal effects. Unobserved confounders (e.g., confidence, risk tolerance) may influence predictors and outcomes, raising the possibility of endogeneity.

- *Intersectional analysis:* Guided by my conceptual framework, I also tested whether rural women face greater challenges and self-employed women experience greater benefits by incorporating interaction terms into the same model structure. In these cases, the *Female* variable in my model structure was replaced with the relevant interaction term. Where significant, I proceeded with further modeling, mirroring the same three-step structure, and additionally estimating a fourth model with separate controls for gender and the subgroup variable to isolate the effect of their intersection. This selective, hypothesis-driven modeling strategy allowed me to test for intersectional differences without overfitting or overcomplicating the analysis, while remaining aligned with the study's theoretical expectations. This is in line with research highlighting that intersectional approaches for theoretically meaningful subgroups can improve validity and relevance of equity-related research (Bauer, 2014).
- *Mediation analysis:* To test theoretically motivated pathways, I conducted mediation analysis using both the Baron & Kenny framework (Baron & Kenny, 1986) and bootstrapped estimation of indirect effects (Preacher & Hayes, 2004). I focused on outcomes where gender was significantly associated with challenges or benefits (e.g., needing help, receiving loan offers), and evaluated whether digital or financial inclusion mediated those effects. Details of the mediation methodology and robustness checks (e.g., Sobel test, bootstrapping) are presented in a dedicated section on mediation before presenting the results.

7.5.2. Computational Text Analysis to Complement Survey Findings

To contextualize and deepen interpretation of the structured survey results, I conducted text analysis on user reviews scraped from the Google Play store for the same fintech app in Kenya. This provided access to organically generated, unstructured user narratives about their experiences with digital loans. I used a combination of computational methods – including sentiment analysis, topic modeling, and word embeddings – to explore patterns in the review data. I used sentiment analysis, an NLP technique, to understand and identify positive and negative reviews. Topic modeling, also a branch of machine learning and NLP, uses statistical language models to uncover latent themes in text. I applied topic modeling specifically to negative reviews to examine recurring challenge-related themes, including how they varied by gender. Finally, I conducted targeted analyses using word embeddings to explore whether key concepts identified in the survey (e.g., needing help) also appeared in the review data and whether their usage differed by gender. This multi-pronged approach allowed for both exploratory analysis of organically emerging themes and comparative analysis aligned with survey findings. I then integrated insights from the review data with the primary survey analysis to enrich and extend interpretation of gendered patterns in technology use and experience.

More detail for each stage and approach is provided below:

- *Data scraping*: I scraped all reviews that were available in January 2024 by importing the required packages from the google_play_scraper import app. The dataset includes reviews provided on a date ranging from April 2020 to August 2024. In total, 234,740 reviews were scraped; of which, 38,344 had names associated with the review. All reviews were in English.
- *Identifying gender of reviewers:* I utilized the Python package, gender-guesser, which was found to have the lowest misclassification rate and appropriate for global data in an academic review (Santamaría & Mihaljević, 2018). The algorithm allowed me to identify gender for a subset of the sample (18,799 reviews) by assessing the 38,344 reviews that included the reviewer's name. A

subset was randomly checked using names and Facebook profiles to ensure that gender was correctly identified. Furthermore, I identified reviews where there is specific reference to gender (e.g., woman, man) were used. This was rare, though several reviewers mentioned being a “business woman”. In each case, the gender assignment was correct with their self-identification.

- *Preprocessing*: I began with standard NLP preprocessing using the SpaCy library, which is a Python library to process text (Vasiliev, 2020). This included removing stop words, removing html links, tokenizing the text, and lemmatizing words to their base forms. I conducted other standard preprocessing such as removing unwanted tokens, and training bigram and trigram models.
- *Sentiment analysis*: VADER is a rule-based model for sentiment analysis (Hutto & Gilbert, 2014), which is popular for its recognition as being more accurate in shorter pieces of text including user reviews (Barik & Misra, 2024). I used the VADER analyzer to assess overall sentiments/polarity (positive, negative, neutral). While various challenges have been identified in sentiment analysis (e.g., related to the presence of grammatical errors and assessing subjective experiences in different language varieties), it remains informative in providing high-level trends around what users or respondents think in short comments and reviews (Wankhade et al., 2022). Overall, I use sentiment analysis here to elucidate high-level perceptions and sentiment by gender.
- *Topic modeling:* I conducted topic modeling using LDA on the negative sentiment reviews. I identified the ideal number of topics by using a coherence model to assess coherence of different topic numbers, and developed visualizations in the form of word clouds and bar charts. LDA is a generative probabilistic model used for topic modeling (Blei et al., 2003), which is popular due to its ability in topic discovery and semantic mining (Jelodar et al., 2019). I used a visualization package (pyLDavis) for each of the topic areas to identify and visualize the most frequent or relevant terms.
- *Word embeddings*: I employed word embedding models (using Word2Vec) – which are dense vector representations of words that capture semantic relationship between words based on

context and co-occurrence (Moreo et al., 2019) – to examine both the frequency and context of terms related to key themes identified in the survey data particularly those with significant gender differences, such as needing help and not receiving loan offers. This included identifying clusters of conceptually related language and using chi-squared tests to assess statistical significance of gender differences.

I also employed computational text analysis in open-ended responses in the survey itself. This was meant to add additional nuance and interpretive depth to the quantitative results. I used similar methods, specifically sentiment analysis and topic modeling, to identify trends and themes within open-ended responses. This approach aligns with recent work emphasizing the value of open-ended survey questions as sites for sociological listening (Newton et al., 2024), where participant voices offer rich insight beyond what is captured in closed-format responses. Illustrative quotes were selected to exemplify themes that emerged across both the quantitative findings and the computational text analysis, and to highlight particularly vivid or nuanced expressions of user experience. Taken together, this strategy enables deeper exploration within the survey data and triangulation across datasets.

7.5.3. Supplementary and exploratory analyses

The initial exploratory findings provide context on the user base and their loan behaviors. These analyses involve comparing financial inclusion indicators among app users with national-level benchmarks to assess if users coming to the app are more or less financially included than the average Kenyan. I also examine loan behavior – such as loan size, late payments, and defaults – using both descriptive statistics and regression analysis disaggregated by gender. For loan size in particular, I go beyond mean comparisons to assess disparities at different points in the distribution through quantile analysis and nonparametric statistical tests (e.g., Mann-Whitney U). While these analyses do not directly test hypotheses, they help contextualize the patterns and inform interpretation of subsequent results.

In the discussion section, I integrate findings from the descriptive and regression results to highlight overall gendered patterns. I return to my hypotheses and weave in components of my conceptual framework to contextualize findings. Together, this analytical approach supports a nuanced understanding of how gender mediates experiences of ML-based credit assessment apps in Kenya.

7.6. Exploratory Findings: Who uses the apps and what are their loan behaviors?

Before presenting the results of hypothesis-driven analysis and turning to my specific research question, I first examine who is using these credit assessment apps and how they interact with loan products. This section explores the demographic profile and financial inclusion status of users in the sample, comparing these characteristics with national benchmarks. It also analyzes users' self-reported purposes for borrowing, as well as key lending behaviors by gender, including loan size, late payment, and default. While the survey sample is not necessarily representative of the broader population using these tools given the potential for self-selection bias, it nevertheless provides exploratory insights. While not directly testing the study's hypotheses, these findings highlight relevant disparities that inform the subsequent analysis.

7.6.1. User demographics

There is a lack of information about who is taking out loans via ML-based alternative lending apps in Kenya, and more broadly. At a high level, the Lender reported that men were more often borrowers, noting approximately 60% of borrowers are men. In the review data scraped from the fintech in Kenya, 39% of the reviewers whose gender I identified were identified as female, while 61% were identified as men. In my effort to have women comprise half of my survey sample, I targeted women in push notifications sent via the Lender and added quotas in Qualtrics.

In the survey, I was able to gain insight into the overall trends of borrower characteristics, including female and male differences, through gathering self-identification information (see Tables 4 and 5 in the

Methodology chapter). Overall, female respondents are more often single (56.25% of females versus 32.42% of men), younger (70% of female respondents are between 18 and 34 years of age compared to 56.6% of men), and report using this fintech app as their sole source of app-based loans (53.75% versus 35.16%, respectively).[21] Male respondents are more often formally employed (46.7% compared to 33.75% of women respondents) with greater proportions of the female sample in informal work than men. This includes higher proportions of female respondents being self-employed (33.12% versus 31.87%), in farming or agriculture (11.25% versus 7.14%), a casual worker (10.62% vs. 6.04%), or having no income either due to being unemployed or a student (6.88% versus 4.95%).

7.6.2. Financial inclusion status of borrowers compared to national averages

The sample revealed that many of the people who take loans out via this app are outside of formal financial folds. Overall, 44.47% of the sample had taken out a formal loan from a financial institution or neobank (52.75% of male respondents, 35.62% of female respondents). Secondly, 54.36% had opened a savings account with a bank or neobank (57.14% of men, 51.88% of women), and only 19.18% had taken out formal insurance (24.73% of men, 13.13% of women).

While this illustrates how many people – particularly women – the apps are reaching who are outside of formal financial folds, I wanted to understand how this compared to Kenya more broadly. I conducted statistical significance tests comparing the survey sample to the broader Kenyan population drawing on data from the World Bank's Global Findex Database (2021). I examined key financial inclusion indicators for which national data exists (i.e. having received a formal loan, having received an informal loan, having opened a formal savings account, and digital financial service usage) to determine whether the sample of app users have greater or lesser (formal) financial access compared to the national average. I used one-sample proportion Z-tests to compare the percentage in each category against the national data and identify p-values. While app-based lending inherently targets smartphone users, this comparison

[21] Overall, both men and women tended to have more than one loan app on their phone. The average number of loan apps for men was reported at 3.2, while for women it was 2.6.

allows for an assessment of how the proportion of app users who are outside formal financial systems compares to the proportion observed in the Kenyan population overall.

Findings indicate that app users tend to be more formally financially included to begin with, as compared to the overall Kenyan population (see Table 7 and Figure 3). I find that 52.74% of men in the sample had taken a formal loan, significantly higher than the 25% national male average, while women in the sample (35.62%) had significantly higher formal loan access compared to the 19% national female average. Both men and women in the sample relied significantly less on informal loans compared to national rates: 35.6% of women in the sample have taken an informal loan and 32.41% of men (compared to 53% of Kenyan women overall and 56% of men). Savings account ownership was notably higher among the sample (54.36%) compared to the national average (21%), reinforcing that these individuals have stronger formal financial access than the general Kenyan population. Still, many of the survey respondents express that they do not have other options outside the app to access a loan for the purpose(s) they needed: 33.52% of male respondents and 42.88% of women report the same. In short, many users of the app lack other loan options, while results suggest higher formal financial inclusion compared to national averages.

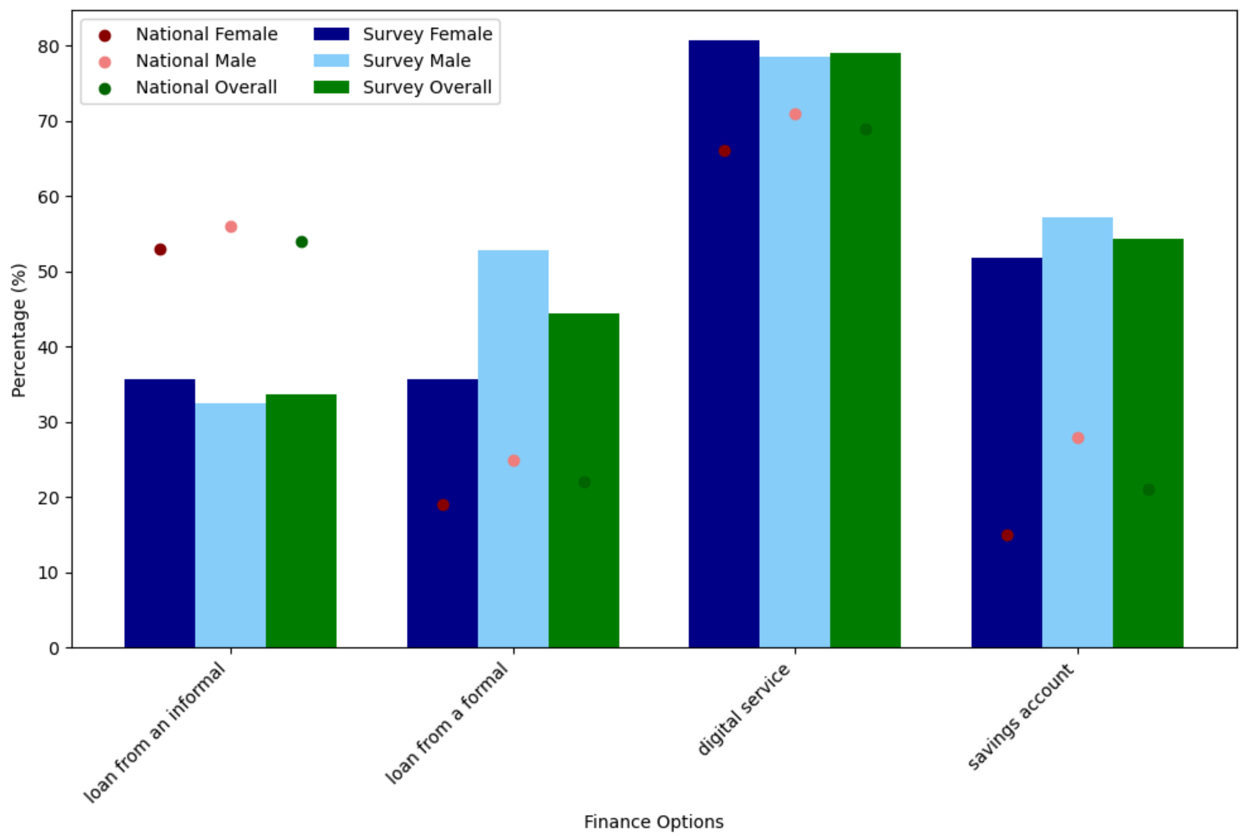
Table 7. Financial inclusion indicators compared to national averages

	Male (%)	**Female (%)**	**Total (%)**	**Overall Kenyan pop (male, %)**	**Overall Kenyan pop (female, %)**	**Overall Kenyan pop (total, %)**
Taken an informal loan	32.41**	35.63**	33.72**	56[A]	53[A]	54[A]
Taken a formal loan from a bank or neobank	52.74**	35.62**	44.47**	25[B]	19[B]	22[B]
Use a digital money service (e.g., M-Pesa)	78.57**	80.63**	79.1**	71[C]	66[C]	69[C]
Opened a savings account with a bank or neobank	57.14**	51.87**	54.36**	28[D]	15[D]	21[D]

**Statistically significant compared to the Kenyan population at $p < 0.05$.
[A] Borrowed from family or friends (age 15+)
[B] Borrowed from a formal financial institution (not including neobanks) (age 15+)

C Have a mobile money account (age 15+)
D Saved at a formal financial institution (age 15+)
All Kenyan overall statistics drawn from (World Bank, 2021).

Figure 3: Financial inclusion indicators by gender, compared to national averages



The sample also shows higher usage of digital money services compared to the average Kenyan population. In particular, 80.63% of the women in the study sample have used another digital money service (compared to 66% of Kenyan women overall) and 78.57% of men have used another digital money service (compared to 71% of Kenyan men overall). This higher percentage among survey respondents makes sense given that the respondents are also using an app for digital loans.

7.6.3. Digital financial inclusion status and app usage of borrowers

To understand the digital financial inclusion status of respondents, I looked at several trends. First, to understand digital inclusion broadly, I used the proxy of whether respondents used their own phone to access the app. Only 12 respondents reported using another's phone (or 3.5% of the sample) reflecting a relatively digitally connected population amongst both men and women. To explore digital financial inclusion related to lending, I asked borrowers if they used any other digital loan apps, and if so, how many. I also asked how many loans they had taken from the particular loan app. Descriptively, women appear less engaged in the digital credit ecosystem, with men using an average of 3.2 digital loan apps compared to 2.6 among women. Regression analysis confirmed that women were significantly more likely to rely exclusively on the Lender's app ($\beta = 0.76$, $p = 0.001$). A broader regression model including demographics and digital financial inclusion predictors of gender provided additional insight: Women in the sample were significantly more likely to be exclusively using the Lender's app ($p = 0.002$) but not more likely to be first-time or initial borrowers. Actually, they were significantly less likely to fall into the 0 to 1 loan category ($p = 0.033$). This finding suggests that women may be more deliberative in their platform use and may reflect more cautious digital loan engagement or selective trust.

I examined the usage of the Lender's app. Of the 342 respondents, the largest share (39.5%) fell into the highest borrowing category (10+ loans), followed by 2 to 5 loans (29.2%) (see Table 8). Gender differences were modest. Among both men and women, the most respondents were in the category of having taken 10+ loans via the app. Men in particular were more concentrated at this higher level: 44.0% were in the 10+ group. Women were more evenly spread across the categories (outside of being first-time borrowers). A chi-squared test of independence indicated that these differences were not statistically significant ($\chi^2(3) = 4.46$, $p = 0.216$), suggesting no strong association between gender and digital borrowing intensity.

Table 8. Loans taken via the lender's app

	Female (%)	**Male (%)**	**Overall (%)**
0-1 Loans	10.0	11.5	10.8
2-5 Loans	31.9	26.9	29.2

6-9 Loans	23.8	17.6	20.5
10+ Loans	34.4	44.0	39.5

7.6.4. Loan purpose

The survey asked respondents about the purpose for their latest loan. While the purpose for taking out the loans varied (see Table 9), both female and male respondents highlighted the use of the loan for day-to-day purposes as the highest prevalence, followed by buying inventory or farming supplies. The third most prevalent loan purpose for women is education, followed by medical needs, and then starting a business. For men, the third most prevalent loan purpose is medical needs, followed by education and then starting a business. These trends align with national Kenyan statistics, whereby women tend to borrow frequently and more than men for education, men tend to borrow more for health or medical purposes, and a third high category for loan purposes for both is to start or expand a business (World Bank, 2021).

Table 9. Survey respondent loan purpose

	Total (%)	**Male (%)**	**Female (%)**
Day to day	34.9	33.15	36.88
Buy inventory or farming supplies	30.79	34.8	26.25
Education	19.06	16.02	22.50
Medical needs	17.3	16.57	18.13
Start a business	14.66	15.47	13.75
Other emergencies	9.59	8.24	11.25
Personal enjoyment	4.69	6.63	2.5
Other	5.28	4.42	6.25
Travel	3.81	5.52	1.88

Another debt	3.81	1.65	6.25
Home improvement	2.35	1.66	3.12
Airtime	1.47	0.62	2.21

Combining some of the categories in the survey data paints a clearer overall picture. When combining all purposes related to business needs and including agricultural activities (i.e. start a business, inventory, farming supplies), the overall percentage is 45.45%, with 40% of women taking out loans related to business needs and 50.27% of men. When combining all purposes related to emergencies (i.e. medical needs, other emergencies), the overall percentage is 27%, with 29.37% of women taking out loans for emergencies and 24.86% of men. This simplified picture of loan purposes reveals that the most common category for taking out a loan is related to business needs, followed by day-to-day needs, and emergencies. This tracks with the Google Play big data analysis, which highlighted two primary use cases for the loans of emergencies and business needs.

7.6.5. Loan size and repayment behavior

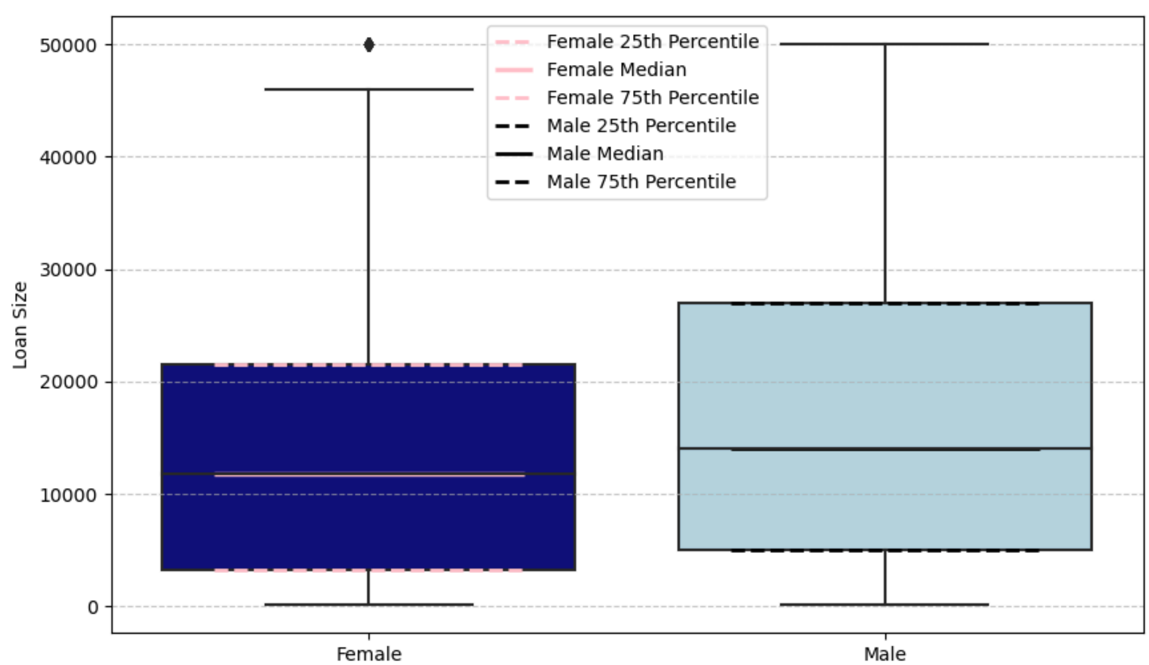
The survey allowed insight into loan and payment behaviors, revealing higher loan sizes for men than women. The analysis shows that the mean loan size for males (18,216.77 Kenyan Shillings (KSh)[22] or 140.78 USD) is higher than for females (15,117.63 KSh or 116.83 USD), with a difference that is marginally significant ($p = 0.0607$). The median loan size for women is 11,750 KSh (90.80 USD) and men is 14,000 KSh (108.19 USD), though not statistically significant ($p = 0.153$).

To further evaluate whether disparities existed at different points in the loan distribution I used quantile analysis and statistical tests to examine gender disparities at the 25th, 50th (median), and 75th percentiles. I calculated loan size quantiles separately for men and women to compare equivalent points in their distributions. I tested differences for significance using the Mann-Whitney U test and compared within a

[22] The exchange rate between Kenyan Shillings and the US Dollar is: USD 1: 129.4 KsH. Retrieved from xe.com on June 8, 2025.

+/- 10% range around each quantile to capture localized differences more accurately. Results showed significant disparities at lower quantiles: women received 1,750 KSh (13.52 USD) less at the 25th percentile ($p = 0.0002$) and 2,250 KSh (17.39 USD) less at the median ($p = 0.0012$). At the 75th percentile, the gap was 5,500 KSh (42.50 USD) but not significant ($p = 0.1742$). See Figure 4 for a visual of loan size distribution including at the different percentiles. Bootstrap confidence intervals for the exact loan size differences at each quantile – which I conducted to measure significance of precise differences in loan sizes – included zero, indicating uncertainty about the precise magnitude of the differences. Overall, these results reveal that women's loans are generally smaller, though the precise dollar amount of the difference at each quantile is uncertain. While this descriptive analysis cannot isolate gender as the sole explanatory factor due to differences across the male and female samples, the trend aligns with qualitative findings from interviews.

Figure 4. Loan size distribution in KSh by gender, with percentiles



Survey results indicated that female borrowers report being late on payments less frequently and report defaulting less often compared to male borrowers. Of female borrowers, 37.74% report late payments versus 45.56% of male borrowers, though the difference was not statistically significant ($p = 0.192$). Both men and women reported similar reasons for being late on payments and for defaulting: income loss, followed by unexpected expenses. The most common reason for late payments for men and women is a lack of money (at a high level of 30.21% of all survey respondents), followed by forgetting the date (4.11%). Similarly, although default rates differ descriptively by gender, the difference is not statistically significant: 11.04% of female borrowers reported defaulting on their loans, compared to 15.08% of male borrowers ($p = 0.3552$).

As a secondary exploratory analysis, I conducted a logistic regression analysis to assess whether user characteristics were associated with the likelihood of late payment and defaulting on a digital loan. Both models included gender (Female), demographic variables, and digital app usage and financial inclusion indicators. For late payment, the model's explanatory power was limited (Pseudo $R^2 = 0.032$), and the overall fit was only marginally significant (LLR $p = 0.097$). However, two predictors were statistically significant: age (younger borrowers were more likely to not have late payments ($p = 0.029$)) and whether the Lender's app was the sole digital loan app (less likely to have late payments ($p = 0.035$)). For default, the overall model fit was modest (Pseudo $R^2 = 0.057$), and the likelihood ratio test was not statistically significant at the 5% level ($p = 0.135$). However, two individual predictors were statistically significant: education (those with a university degree are significantly less likely to default ($p = 0.043$)) and whether the Lender's app was the sole digital loan app (those with only the app were significantly less likely to default ($p = 0.024$)). Gender was not statistically significant in either of the models. Therefore, the gendered patterns observed in late payment and default may be more driven by correlated factors, whereby female borrowers tend to be younger and only have the Lender's app for digital loans.

7.7. Main findings: Challenges, benefits, and pathways

This section presents both descriptive and inferential findings. Rather than separating results by method, I organize them thematically into two areas: challenges users experience and benefits they perceive. This structure clearly illustrates how gender shapes users' experiences with ML-based alternative lending apps. As these analyses are based on cross-sectional, observational data, the findings reflect associations rather than causal relationships.

7.7.1. Challenges experienced

7.7.1.1 Downloading and using the app

Overall, respondents reported that downloading and using the app was fairly easy. However, women report marginally higher levels of challenge than men in downloading and using the app. On a scale of 1 (very easy) to 5 (very challenging), descriptive statistics show that women reported slightly higher average challenge levels in downloading and using the app (M = 1.35, SD = 0.73, 95% CI [1.24, 1.46]) compared to men (M = 1.21, SD = 0.57, 95% CI [1.13, 1.30]). A t-test indicated that this difference was marginally significant ($t = -1.90$, $p = 0.058$).

To better examine the relationship between gender and challenges in downloading and using the app, I conducted a series of ordered logistic regressions. In the baseline model, being female was associated with greater reported difficulty, though the effect was only marginally significant (β = 0.52, OR = 1.69, 95% CI [0.98, 2.92], $p = 0.062$). In Model 2, after controlling for demographic characteristics, the coefficient for gender was reduced and became statistically non-significant (β = 0.41, $p = 0.166$). In Model 3, which added financial inclusion and digital app usage variables, the gender effect remained non-significant (β = 0.44, $p = 0.161$), while formal financial inclusion emerged as a strong and significant predictor of not facing challenges (β = –1.18, $p < 0.001$) (see Table 10). This suggests that exclusion from the formal financial system plays a critical role in shaping users' ease with digital tools. Other factors such as low educational attainment and rural residence were not significant in the full model, indicating that these demographic traits alone do not explain the gender gap in reported app-related challenges.

Table 10. Logistic regression results: Challenges in downloading and using the app

Variable	Model 1: Baseline Coef (SE)	Model 2: Demographics only Coef (SE)	Model 3: Full controls Coef (SE)
Female	0.52 (0.28)*	0.41 (0.30)	0.44 (0.31)
Married	–	-0.61 (0.32)*	-0.53 (0.33)
Employed	–	0.37 (0.30)	0.64 (0.31)
Youth (18-34 years old)	–	0.08 (0.32)	-0.27 (0.35)
Rural	–	0.14 (0.40)	0.15 (0.41)
University degree	–	-0.72 (0.29)**	-0.48 (0.31)
Lender's app only digital loan app used	–	–	-0.36 (0.31)
0 - 1 loans taken	–	–	0.64 (0.43)
Financial inclusion	–	–	-1.18 (0.32)***
Threshold 1|2	1.75 (0.21)***	1.28 (0.41)***	0.43 (0.50)
Threshold 2|3	0.02 (0.16)	0.04 (0.16)	0.09 (0.16)
Threshold 3|4	0.43 (0.24)*	0.44 (0.24)*	0.46 (0.23)**
Model fit (LogLik / AIC)	-218.36 / 444.7	-212.92 / 443.8	-204.40 / 432.8
Observations (N)	342	342	342

Notes: Coefficients represent log-odds from ordered logistic regression predicting greater reported challenges.
Standard errors are shown in parentheses. p-values are based on two-tailed tests. *p < 0.1, **p < 0.05, ***p < 0.01.
All VIFs for independent variables were below 2.5, indicating no problematic multicollinearity.

There are notable gender differences in reported need for help using the app: 12.5% of women reported needing help, compared to just 6.0% of men. In a baseline logistic regression, women were significantly less likely than men to report *not* needing help (β = –0.80, OR = 0.45, 95% CI [0.21, 0.97], p = 0.042). This odds ratio indicates that women's odds of reporting not needing help were 55% lower than men's, corresponding to nearly double the likelihood of women reporting a need for help. In a second model adjusting for demographic characteristics, the gender effect was reduced and no longer statistically significant (β = –0.51, OR = 0.60, 95% CI [0.27, 1.35], p = 0.216), suggesting that some of the observed gender gap may be partly explained by demographic differences. In the full model, including controls for digital loan app usage and formal financial inclusion, the gender coefficient remained non-significant (β = –0.46, OR = 0.63, 95% CI [0.27, 1.45], p = 0.276). Notably, financial inclusion emerged as a marginally significant predictor (β = 0.78, OR = 2.19, 95% CI [0.96, 4.99], p = 0.063 (see Table 11)). Together, these results suggest that while women initially appear more likely to need assistance using the app, this association is attenuated when accounting for structural factors like education and financial inclusion.

Table 11. Logistic regression results: Help not needed, model comparison

Variable	**Model 1: Baseline Coef (SE)**	**Model 2: + Demographics Coef (SE)**	**Model 3: + Financial use / digital inclusion Coef (SE)**
Intercept	2.74 (0.31)***	2.57 (0.63)***	2.40 (0.74)***
Female	-0.80 (0.39)**	-0.51 (0.41)	-0.46 (0.43)
Married	–	0.94 (0.46)**	0.80 (0.48)*
Employed	–	-0.07 (0.43)	-0.23 (0.45)
Youth (18-34 years old)	–	-0.94 (0.53)*	-0.67 (0.55)
Rural	–	-0.30 (0.55)	-0.25 (0.56)
University degree	–	0.89 (0.41)**	0.64 (0.43)
Lender’s app only digital loan app used	–	–	-0.33 (0.42)
0 - 1 loans taken	–	–	-0.74 (0.52)
Financial inclusion	–	–	0.78 (0.42)*

Model Fit (LogLik / Pseudo R^2)	-101.81 / 0.021	-93.95 / 0.096	-90.90 / 0.126
Observations (N)	342	342	342

Notes: Coefficients represent changes in the log-odds of the outcome (logistic regression).
Standard errors (SE) in parentheses. *p*-values from two-tailed tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
All VIFs for independent variables were below 2.5, indicating no problematic multicollinearity.

Women expressing need for help downloading and using the app is echoed in the Google Play review data. While both men and women discuss challenges related to access (downloading the app, applying to the app, and accessing customer service), use (paying back loans and not understanding terms), and uncertainty (being denied and not understanding why), there are some slight gender differences. Using text analysis and NLP methods discussed above, I identified occurrences of "help" in user reviews and compared the frequency between men and women. Overall, 8.22% of female users mentioned "help", compared to 7.20% of male users, found statistically significant through a Chi-Square test ($\chi^2 = 6.48$, $p = 0.011$). This pattern is consistent with the descriptive survey results, though regression analysis in the survey results shows that the gender difference is attenuated once demographic and financial inclusion factors are taken into account.

7.7.1.2 Particular challenges in using the app, accessing money, and repayment

The survey also asked about specific challenges, including whether the app wouldn't open or had glitches that made it difficult to use; not receiving a credit or loan offer; receiving a credit / loan offer, but never received the loan monies in one's bank account; not receiving reminders for my loan payments; and poor customer service in responding to queries or asks for assistance. The most common challenge, for both men and women, is that the app wouldn't open or had glitches that made it difficult to use: 11.44% of respondents report this as a challenge, followed by inability to obtain a loan (6.45%) and not receiving money after loan approval (5.28%). I conducted a Chi-square test to assess whether the distribution of challenges differed between male and female borrowers. The results indicated that none of the reported challenges showed statistically significant differences between genders ($p > 0.05$) (see Table 12).

Table 12. Challenges experienced, overall and by gender

Challenge	Overall (%)	Female (%)	Male (%)	P-value (Gender)
App glitches	11.2%	11.9%	11.5%	1.0000
No credit or loan offer received	6.2%	7.5%	5.5%	0.5937
Money not received	5.0%	6.2%	4.4%	0.6005
No repayment reminders received	4.2%	3.8%	4.9%	0.7842
Poor customer service	4.5%	3.1%	6.0%	0.3083
Other problems	1.7%	1.9%	1.6%	1.0000

Chi-square tests assess whether reporting rates differ significantly by gender.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

While gender differences were not statistically significant, descriptive patterns emerged in the types of challenges reported. Women more often reported challenges related to accessing money and the app (having the app glitch, not receiving a loan, and/or not receiving money even if approved for a loan). Meanwhile, men more often reported challenges related to repayment and customer service (not getting reminders for payments, and poor customer service in responding to queries or asks for assistance). These patterns are descriptive in nature, small in magnitude, and are not interpreted as statistically meaningful gender differences.

The Google Play reviews analysis reflects the data. I conducted topic modeling for negative sentiment reviews (n = 2,582) (see Appendix 5, Figure 1). While there are largely similarities between men and women, some differences emerge: Women tend to discuss loan eligibility, payment processing, and financial experiences. Men focus more on service interactions and technical issues. Both groups share common concerns regarding financial approvals, denials, emergency needs, and frustration with payment issues. Similarly to the descriptive survey findings, these findings suggest women are more focused on procedural and eligibility aspects, whereas men highlight service quality and technical issues.

7.7.1.3 Not receiving a loan offer

For each challenge, I ran baseline logistic regressions to test whether gender was significantly associated with outcomes. I did the same with the interaction term, rural women, in line with the hypothesis that this subgroup may face greater challenges. Only one challenge yielded a statistically significant result: rural women and the challenge of not receiving a loan offer. The other challenges were not significantly related to gender differences overall or at this interaction.

In a baseline logistic regression (including only an intercept and an interaction term for being both female and living in a rural area), rural women had 5.24 times higher odds of not receiving a loan offer ($\beta = 1.66$, OR = 5.24, 95% CI [1.73, 15.91], $p = 0.003$). The intercept for this model was negative ($\beta = -2.88$, $p < 0.001$), indicating a low baseline probability of this challenge for the reference group (non-rural women and men). These findings align with descriptive statistics: while 6.4% of all respondents reported this challenge (7.5% of women and 5.5% of men), the rate rose to 22.7% among rural women ($\chi^2 = 8.49$, $p = 0.014$). A demographics model including the rural-female interaction and controls showed that rural women remained significantly more likely to report this challenge ($\beta = 1.46$, OR = 4.31, $p < 0.05$). This association is maintained and slightly strengthened in a model that additionally includes financial inclusion and digital loan use variables ($\beta = 1.58$, OR = 4.86, $p < 0.01$). In the fully adjusted model that also includes the main effects of gender and rural residence, the interaction term is no longer statistically significant (see Table 13). This suggests that the increased likelihood among rural women is not explained by differences in observed demographic or financial inclusion characteristics, but instead reflects the combined, additive effects of gender and rural residence.

Multicollinearity was ruled out (VIFs for female, rural, and rural female were all under 2.1), indicating the reduced significance is not due to multicollinearity but overlapping explanatory factors. Stratified models further support this: rural residence was a significant predictor for women (β = 1.27, OR = 3.55, p = 0.014), but not for men (β = 0.61, OR = 1.83, p = 0.454). Because men and women have nearly identical rural representation in the sample (13.81% of men and 13.84% of women), this pattern is not driven by compositional differences. This suggests that rural context is more strongly associated with barriers to credit access for women.

Table 13. Logistic regression results: Rural women not receiving a loan offer

Variable	Model 1: Baseline Coef (SE)	Model 2: + Demographics Coef (SE)	Model 3: + Financial use / digital inclusion Coef (SE)	Model 4: Full controls Coef (SE)
Intercept	-2.88 (0.25)***	-2.39 (0.57)***	-1.87 (0.71)**	-1.83 (0.78)**
Rural female	1.66 (0.57)**	1.46 (0.59)**	1.58 (0.61)***	1.03 (1.09)
Female	–	–	–	-0.18 (0.56)
Rural	–	–	–	0.69 (0.86)
Married	–	-0.71 (0.51)	-0.61 (0.52)	-0.72 (0.55)
Employed	–	-0.32 (0.52)	-0.41 (0.54)	-0.40 (0.54)
Youth (aged 18-34)	–	0.09 (0.51)	-0.06 (0.54)	-0.04 (0.55)
University degree	–	-0.23 (0.46)	-0.14 (0.48)	-0.10 (0.49)
Lender’s app only digital loan app used	–	–	-0.56 (0.49)	-0.53 (0.50)
0 - 1 loans taken	–	–	0.46 (0.63)	0.47 (0.64)
Financial inclusion	–	–	-0.56 (0.50)	-0.57 (0.51)
Model fit (Log likelihood / Pseudo r-squared)	-76.40 / 0.064	-76.40 / 0.064	-74.96 / 0.082	-74.53 / 0.087
Observations (N)	342	342	342	342

Notes: Coefficients represent log-odds of logistic regression model.

Standard errors (SE) in parentheses. *p*-values from two-tailed tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
All VIFs for independent variables were below 2.5, indicating no problematic multicollinearity.

Descriptive patterns reinforce this interpretation. Rural women in the sample were less often formally employed compared to the overall sample average (22.7% vs. 40.64%), more likely to be self-employed or working in agriculture, and less likely to hold a university degree. Yet, their repayment performance was not worse: their default and late payment rates did not differ significantly from the broader sample[23]. This challenges assumptions that rural women represent greater credit risk.

The findings of women experiencing the challenge of not receiving a loan offer (or being denied) are echoed in the Google Play review data results. Female users discuss being denied a loan slightly more than men. I used a word embedding model to identify terms semantically similar to "denial" (i.e., words with similar vector representations in the dataset). From this, I compiled the list of denial-related terms (e.g., "denied", "rejected", "declined") and analyzed their frequency across male and female user reviews. I conducted a Chi-Square test to determine whether these differences in mention rates were statistically significant. Women were significantly more likely than men to mention being denied a loan in their app reviews. Specifically, 1.12% of female users mentioned denial-related terms, compared to 0.81% of male users. A Chi-Square test confirmed that this difference was statistically significant ($\chi^2 = 4.38$, $p = 0.036$), representing roughly a 38% higher likelihood for women. Although the absolute share of reviews mentioning denial-related terms is small, women are more likely than men to reference them. This may reflect greater salience of loan denial among female users or may reflect higher underlying denial rates, though review data cannot distinguish between these explanations. In examining "default", men referenced default related terms slightly more often than women (0.76% vs. 0.67%), however a Chi-Square test found this difference not to be statistically significant ($\chi^2 = 0.36$, $p = 0.549$).

[23] Rural women had slightly lower late payment rates (36.4% vs. 41.5%) and slightly higher default (13.6% vs. 12.9%) compared to the overall sample. The difference in late payment rates was not statistically significant ($\chi^2 = 0.08$, $p = 0.777$), nor was default ($\chi^2 = 0.00$, $p = 1.00$).

The user reviews help bring these trends to life and elucidate more of the emotions and tensions users experience. As one woman said, “This app has just broken my heart right now, I really was relying on it to get some loan to start a business but am disappointed. To make matters worse they have let me fill in the application form only to tell me that I can't be given the loan I needed.” The sense of frustration in being denied a loan is shared by men. A male user noted, “You call it instant loans yet you have the guts to tell someone to wait for 30 days to be eligible again....do you really think some emergencies can wait even for 3 hours? I'm so disappointed in this app...it's very unreliable at a time of emergency.” These quotes illustrate how valuable accessing loans is to people and how the emotional experiences of denial.

7.7.2. Understandings of the loan process

To better understand potential challenges, the survey assessed borrowers’ understanding of various aspects of the loan process, including approval process, terms and conditions, payment process, and usage of personal data. Overall, respondents reported high levels of understanding, with all mean scores above 4.0 on a scale of 1 to 5 (where 1 = completely disagree, and 5 = completely agree). The highest level of understanding was reported for "No trouble understanding payments" (Mean = 4.61, SD = 0.93, $t = 31.80$, $p < 0.0001$), followed by "Understand terms and conditions" (Mean = 4.48, SD = 0.92, $t = 28.77$, $p < 0.0001$) and "Understand the approval process" (Mean = 4.45, SD = 1.07, $t = 24.86$, $p < 0.0001$). The lowest but still positive level of agreement was reported for “Understand how the app uses my data” (Mean = 4.10, SD = 1.18, $t = 20.86$, $p < 0.0001$). When comparing male and female respondents, no statistically significant gender differences were observed for any understanding measure ($p > 0.05$) and mean scores were relatively close across both male and female respondents (see Appendix 6, Table 1). I also explored trust in the Lender. Trust was also reported to be fairly high in the Lender with an overall mean of 4.23 (SD = 1.09, $t = 20.86$, $p < 0.0001$). While men report slightly higher levels of trust than women (Mean = 4.28 versus 4.17), gender differences are not statistically significant.

The positive understanding is echoed in the open-ended survey responses; however, both men and women reveal uncertainty and experience issues. Concerns and uncertainty particularly arise related to not understanding costs of loans and personal data usage. One female respondent said, “It was easy to use though once there was a delay in disbursing the money to my account when I was in need. I asked for a cancellation since I got money somewhere else but they refused.” Another female respondent reported, “I didn't clearly understand the cost of the loan (interest and penalties). About my personal data usage I have no idea how and what the App intends to use my data for.” One male respondent echoed the lack of understanding on data usage saying, “I don't understand how [Fintech] stores and safeguards my data.” The comments on not understanding how data is used reflects how this area of understanding was rated the lowest in the overall sample.

7.7.3. Impacts experienced

7.7.3.1. Impacts overall

The survey prompted respondents to think about the latest loan they took out from the fintech, and then report on how the loan impacted their self-confidence, hopefulness about the future, sense of financial control, ability to make financial decisions, ability to meet basic needs, overall financial well-being, and involvement in large household purchases. Respondents answered questions about impacts experienced on a scale of 1 to 5 (1 = very negatively, reduced significantly; 2 = slightly negative, reduced somewhat; 3 = no change; 4 = slightly positively, increased somewhat; 5 = very positively, increased significantly).

Both men and women report positive impacts. The highest-rated impact was "Hopefulness about the future" (Mean = 4.43, SD = 0.91, $t = 26.75$, $p < 0.0001$) (see Table 15). As one female borrower noted, “It gave me a sense of security and hope when I almost lost hope to live.” The second-highest impact was “Ability to make financial decisions” (Mean = 4.36, SD = 0.89, $t = 28.22$, $p < 0.0001$) and “Self-confidence” (Mean = 4.32, SD = 0.91, $t = 26.75$, $p < 0.0001$). As one male respondent said, “I’m so confident and self-assured when I get the loan.” Another respondent highlighted how the app helps bring a sense of financial control, which was also highly rated across the sample: “I am able to control my finances better. I don’t have to strain thinking that a loan is a bad thing. It has helped meet my stock intake for a business that is still growing.”

One of the lowest-rated but still overall positive impacts was “Overall financial well-being” (Mean = 4.1, SD = 1.02, t = 19.94, $p < 0.0001$). Various borrowers discussed that the app helped them learn to be better financial managers, including helping them save and make future plans, enhance financial decision making, and learn more about financial budgeting. Others discussed help in reaching financial goals and doing financial planning.

While minor differences in means were observed using t-tests, none of the gender comparisons were statistically significant ($p > 0.05$), indicating that both men and women perceived similar benefits from the loans. Additionally, effect sizes (Cohen’s d) were consistently small ($d < 0.20$), reinforcing that gender differences were negligible in the descriptive statistics.

Table 15. Impacts experienced, overall and by gender

Impact	Overall (Mean [SD])	Female (Mean [SD])	Male (Mean [SD])	p (Gender)	Sig (Gender)	p (vs 3.0)	Sig (vs 3.0)
Self-confidence	4.32 [0.91]	4.33 [0.91]	4.32 [1.26]	0.9190	No	0.0000	***
Hopefulness about the future	4.43 [0.85]	4.51 [0.85]	4.36 [0.89]	0.1342	No	0.0000	***
Sense of control over spending money	4.24 [0.96]	4.26 [1.03]	4.23 [0.90]	0.7779	No	0.0000	***
Financial decision-making ability	4.36 [0.89]	4.36 [0.97]	4.36 [0.82]	0.9766	No	0.0000	***
Ability to meet basic needs	4.24 [0.95]	4.29 [0.97]	4.19 [0.94]	0.3636	No	0.0000	***

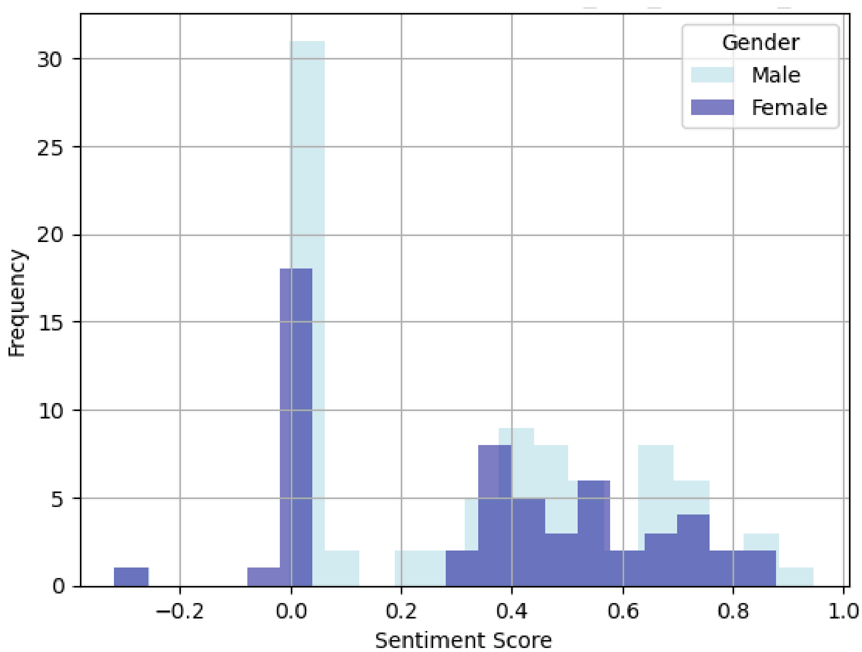
Financial well-being	4.10 [1.02]	4.07 [1.08]	4.13 [0.97]	0.6008	No	0.0000	***
Involvement in large purchases	4.10 [1.01]	4.09 [1.08]	4.11 [0.95]	0.8358	No	0.0000	***

One-sample *t*-tests compare group means to a neutral value of 3.0 on a 1 to 5 scale for overall significance.
Chi-square tests assess whether reporting rates differ significantly by gender.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Sentiment analysis on open-ended responses about impacts experienced reveals moderately positive sentiment overall (average of 0.33[24]) and by gender (0.34 for women, 0.32 for men), with very little negative sentiment (see Figure 5). Several respondents referred to the Lender as a "friend" or "partner". As one male respondent noted, "[Lender] has been good to me coz it's a friend to me in time of need." Various respondents highlight the emotional benefit of the app, including a sense of "calmness" knowing the app was available and supporting a sense of dignity. A male borrower highlighted, "The loan helped preserve my dignity and be able to pay for the hospital bill. Now I will repay my loan without anyone thinking of me as a person who does not [have an] emergency fund set aside." These comments highlight the psychological and emotional benefits felt by users. They also reflect the high use of these loans for emergencies, as discussed in the exploratory findings section.

[24] Sentiment is rated on a scale from -1, most negative, to 1, most positive; with 0 as neutral

Figure 5. Sentiment analysis of open-ended responses about impacts experienced, by gender



Many people taking the loans are also doing so for business-related purposes, as aforementioned (40% of women taking out loans related to business needs and 50.27% of men). Topic modeling highlights the main impacts experienced for those who borrow loans for business purposes: (1) stock management and profitability; (2) stock availability and sales growth; and (3) business expansion and inventory growth (see figures in Appendix 5). As one male respondent noted, "I am able to maintain cyclical stock flow." A female respondent similarly noted, "The loan actually helped me buy some stock for my business which was profitable." While another said, "I was able to pay for farm inputs. I will also be able to increase my income."

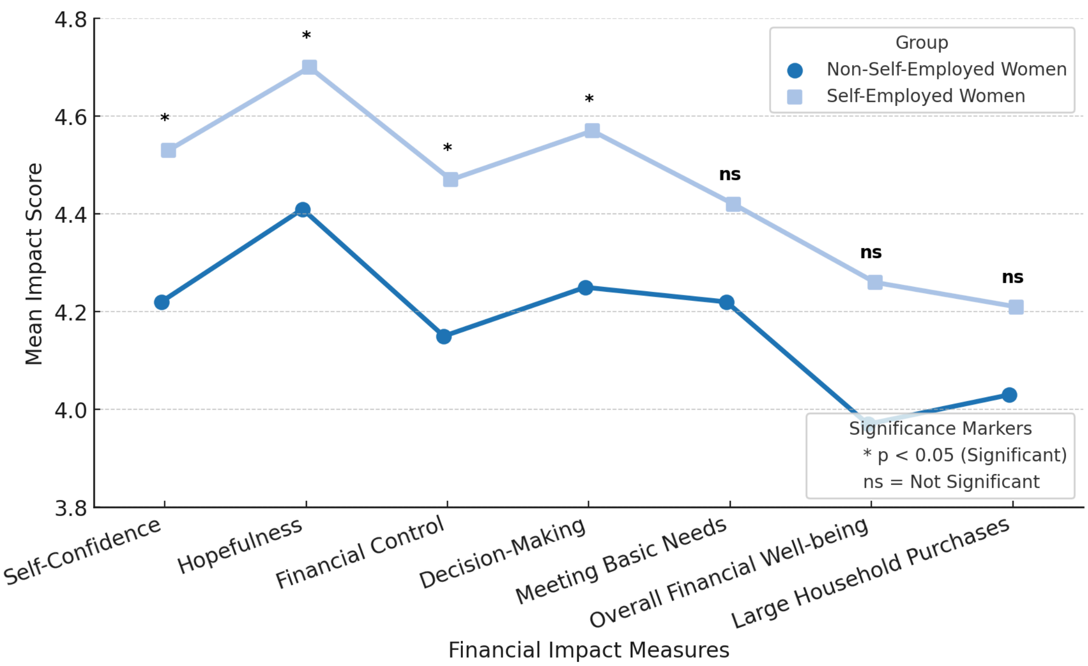
Descriptive statistics show that not all impacts from accessing the app and loans were positive. In particular, 9.7% of respondents reported a decrease in their financial well-being, this was greater for women (11.25%) versus men (8.29%), though this gender difference is not statistically significant. Respondents note unintended consequences, including related to growing reliance on the Lender. One female respondent put it: "I feel that in [the] event of an emergency I may have access to funds through

[Lender's] app and this will help sort the issue, while on the downside I may also end up doing impulse buying because of the same." Other respondents echo downsides related to spending habits and high interest rates. One male borrower said, "The loan enables me to achieve my goal, however I felt like spending more due to the daily interest for the loan. I wish I could be in a position not to continue doing my projects through a loan… The cost of living and shrinking income has caused [my] family to [rely on] lending for consumption." This respondent connects their reliance on the app to the difficult financial situation in Kenya for many people and low levels of financial health. Relatedly, a female borrower spoke to potential concerns of debt traps and ongoing borrowing: "I feel like I am starting to have a dependency on the app." Ultimately, it provides benefits while highlighting how underlying issues in financial distress can remain for users.

7.7.3.1. Impacts for self-employed women

While gender alone does not significantly influence perceived loan impacts, self-employment status among women does (see Figure 6). Self-employed women reported significantly higher mean impact scores than non-self-employed women on self-confidence (Mean = 4.53, SD = 0.72 vs. 4.22, SD = 0.98, $p = 0.0288$, $d = 0.34$), hopefulness about the future (4.70, SD = 0.70 vs. 4.41, SD = 0.90, $p = 0.0282$, $d = 0.34$), sense of financial control (4.47, SD = 0.87 vs. 4.15, SD = 1.09, $p = 0.0449$, $d = 0.32$), and ability to make financial decisions (4.57, SD = 0.84 vs. 4.25, SD = 1.01, $p = 0.0406$, $d = 0.33$). These differences were statistically significant ($p < 0.05$) and had moderate effect sizes (Cohen's $d \sim 0.32$–0.34), suggesting meaningful practical differences in these financial impact measures. Meanwhile, similar tests between self-employed men and non-self-employed men revealed that differences were not statistically significant. These findings suggest that female entrepreneurs experience greater perceived benefits from loans compared to other women, particularly in terms of self-confidence, hopefulness, financial control, and financial decision-making.

Figure 6. Self-employed women experience greater impacts



To further understand gender differences in perceived loan impacts, I conducted ordinal logistic regression models[25] to control for various factors and incorporate intersectionality through interaction variables. I first examined "hopefulness about the future" as this variable had gender differences that were marginally significant in the descriptive statistics. I wanted to understand this for both female overall and female entrepreneurs, 96.2% of whom had reported a positive change in hopefulness about the future, compared to 84.1% of other women and men (p = 0.033).

In a baseline ordered logistic regression model, being female was marginally associated with greater reported hopefulness from accessing and using the app (β = 0.40, OR = 1.50, 95% CI [0.97, 2.31], p = 0.067). Being a self-employed woman was a statistically significant predictor of increased hope (β = 0.89, OR = 2.44, 95% CI [1.24, 4.81], p = 0.010). Self-employed women had 2.44 times higher odds than other

[25] An ordinal model is appropriate given the ordinal nature of the data. A 1-5 scale reflects ordered categories, where distances are not uniform between the different levels (ordinal thereby accounts for this structure). Meanwhile, most responses are clustered at 4 and 5, thus treating the scale as continuous (e.g., through a linear regression) could distort analysis.

women and men of reporting increased hopefulness due to accessing the app and receiving their latest loan.

In an adjusted ordered logistic regression model including demographic controls plus digital loan and financial inclusion variables the association between being a self-employed woman and reporting increased hope remained significant (β = 0.95, OR = 2.60, 95% CI [1.30, 5.21], *p* = 0.007). In an alternative specification that added the main effects of gender and employment to avoid overlap with the interaction term, the effect of being a self-employed became marginally significant (β = 0.80, OR = 2.23, 95% CI [1.00, 4.98], *p* = 0.051) (See Table 16).

Table 16. Logistic regression results: Impact on hopefulness for self-employed women, across models

Variable	**Model 1: Baseline Coef (SE)**	**Model 2: + Demographics Coef (SE)**	**Model 3: + Financial use / digital inclusion Coef (SE)**	**Model 4: Full controls Coef (SE)**
Self-employed female	0.89 (0.35)***	0.91 (0.35)***	0.95 (0.36)***	0.80 (0.41)*
Female	–	–	–	0.18 (0.26)
Employed	–	–	–	-0.11 (0.26)
Married	–	0.11 (0.23)	0.03 (0.24)	0.09 (0.26)
Youth (aged 18-34)	–	-0.10 (0.24)	0.03 (0.26)	0.01 (0.26)
Rural	–	0.13 (0.33)	0.11 (0.33)	0.10 (0.33)
University degree	–	0.06 (0.22)	-0.02 (0.23)	-0.03 (0.24)
Lender's app only digital loan app used	–	–	0.70 (0.24)***	0.65 (0.25)***
0 - 1 loans taken	–	–	-0.46 (0.36)	-0.45 (0.37)
Financial inclusion	–	–	0.55 (0.26)**	0.56 (0.24)**
Has other loan options	–	–	0.30 (0.24)	0.31 (0.24)

Threshold 1|2	-4.11(0.45)***	-4.06 (0.52)***	-3.34 (0.58)***	-3.32 (0.60)***
Threshold 2|3	-0.02 (0.363)	-0.02 (0.36)	-0.013 (0.36)	-0.013 (0.36)
Threshold 3|4	0.34 (0.18)**	0.34 (0.18)*	0.35 (0.18)**	0.35 (0.18**
Threshold 4|5	0.30 (0.10)***	0.31 (0.10)**	0.34 (0.10)**	0.34 (0.10)**
Model fit (LogLik / AIC / BIC)	-343.38 / 696.8 / 715.9	-342.93 / 703.9 / 738.4	-335.46 / 696.9 / 746.7	-335.17 / 700.3 / 757.8
Observations (N)[26]	341	341	341	341

Notes: Coefficients represent log-odds of ordered model.
Standard errors (SE) in parentheses. *p*-values from two-tailed tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
All VIFs for independent variables were below 2.5, indicating no problematic multicollinearity.

I explored whether female entrepreneurs experienced significant differences in other impacts as well through conducting baseline ordered logistic regressions for each impact variable. Two impact variables stood out, the first of which being financial decision-making ability. In the baseline model, women who were self-employed were significantly more likely to report greater positive impacts on their ability to make financial decisions ($\beta = 0.74$, OR = 2.10, 95% CI [1.12, 3.96], $p = 0.022$). This effect remained statistically significant after adjusting for demographic variables, as well as after adding in digital loan usage and financial inclusion variables. However, it became marginal when additional controls for gender and employment were included ($\beta = 0.67$, OR = 1.95, 95% CI [0.92, 4.13], $p = 0.084$) (See Table 17), suggesting some shared explanatory power across these variables.

Table 17. Logistic regression results: Impact on financial decision-making for self-employed women, across models

Variable	Model 1: Baseline Coef (SE)	Model 2: + Demographics Coef (SE)	Model 3: + Financial use / digital inclusion Coef (SE)	Model 4: Full controls Coef (SE)
Self-employed female	0.74 (0.32)**	0.75 (0.33)**	0.70 (0.33)**	0.66 (0.39)*

[26] Observations for analysis on impacts are 341. This is because one respondent did not answer this portion of the survey.

Female	–	–	–	-0.09 (0.25)
Employed	–	–	–	-0.22 (0.25)
Married	–	-0.09 (0.23)	-0.14 (0.23)	-0.12 (0.25)
Youth (aged 18-34)	–	-0.42 (0.24)*	-0.38 (0.25)	-0.41 (0.25)*
Rural	–	0.47 (0.33)	0.45 (0.34)	0.43 (0.34)
University degree	–	-0.00 (0.21)	-0.04 (0.22)*	-0.01 (0.23)
Lender's app only digital loan app used	–	–	0.39 (0.22)*	0.37 (0.23)
0 - 1 loans taken	–	–	-0.34 (0.34)	-0.40 (0.34)
Financial inclusion	–	–	0.34 (0.25)	0.33 (0.25)
Has other loan options	–	–	-0.17 (0.23)	-0.15 (0.23)
Threshold 1|2	-4.12(0.45)***	-4.40 (0.52)***	-4.12 (0.57)***	-4.35 (0.60)***
Threshold 2|3	-0.02 (0.36)	-0.02 (0.36)	-0.02 (0.36)	-0.02 (0.36)
Threshold 3|4	0.45 (0.16)**	0.45 (0.16)**	0.45 (0.16)**	0.45 (0.16)**
Threshold 4|5	0.34 (0.10)***	0.36 (0.10)***	0.37 (0.10)***	0.37 (0.10)***
Model fit (LogLik / AIC / BIC)	-365.38 / 740.8 / 759.9	-362.41 / 742.8 / 777.3	-359.46 / 744.9 / 794.7	-365.33 / 762.7 / 824.0
Observations (N)	341	341	341	341

Notes: Coefficients represent log-odds of ordered model.
Standard errors (SE) in parentheses. *p*-values from two-tailed tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
All VIFs for independent variables were below 2.5, indicating no problematic multicollinearity.

The other impact variable that was significant for female entrepreneurs was one's sense of control over finances. In the baseline ordered logistic regression model, women who were self-employed were significantly more likely to report greater perceived improvements in their control over money as a result of their most recent loan and use of the lender's app (β = 0.64, OR = 1.89, 95% CI [1.05, 3.41], p = 0.035). This effect remained significant after adjusting for demographic variables (β = 0.62, OR = 1.86, 95% CI [1.02, 3.37], p = 0.042). It became marginally significant when adding digital loan usage and financial inclusion variables (β = 0.54, OR = 1.71, 95% CI [0.94, 3.12], p = 0.081). It was not statistically significant when additional controls for gender and employment were included (β = 0.42, OR = 1.53, 95% CI [0.75, 3.14], p = 0.247) (See Table 18), suggesting that part of the observed effect may be explained by broader differences in gender and employment status, which is linked to reasons for taking out a loan (e.g., for business).

Table 18. Logistic regression results: Impact on sense of financial control for self-employed women, across three models

Variable	Model 1: Baseline Coef (SE)	Model 2: + Demographics Coef (SE)	Model 3: + Financial use / digital inclusion Coef (SE)	Model 4: Full controls Coef (SE)
Self-employed female	0.65 (0.30)**	0.62 (0.30)**	0.53 (0.31)*	0.43 (0.34)
Female	–	–	–	-0.10 (0.25)
Employed	–	–	–	-0.34 (0.25)
Married	–	-0.17 (0.22)	-0.14 (0.23)	-0.11 (0.24)
Youth (aged 18-34)	–	-0.18 (0.23)	-0.23 (0.24)	-0.28 (0.24)
Rural	–	0.73 (0.32)**	0.69 (0.33)*	0.68 (0.33)**
University degree	–	-0.09 (0.21)	-0.03 (0.22)	0.004 (0.22)
Lender's app only digital loan app used	–	–	0.51 (0.22)**	0.46 (0.23)**
0 - 1 loans taken	–	–	-0.07 (0.35)	-0.17 (0.35)
Financial inclusion	–	–	0.05 (0.24)	0.03 (0.24)
Has other loan options	–	–	-0.38 (0.22)*	-0.35 (0.22)

Threshold 1|2	-3.95 (0.41)***	-4.13 (0.49)***	-4.16 (0.54)***	-4.42 (0.57)***
Threshold 2|3	0.26 (0.27)	0.26 (0.27)	0.27 (0.27)	0.27 (0.27)
Threshold 3|4	0.26 (0.15)*	0.27 (0.15)*	0.28 (0.15)*	0.28 (0.15)*
Threshold 4|5	0.33 (0.09)***	0.34 (0.09)***	0.37 (0.09)***	0.37 (0.09)***
Model Fit (LogL / AIC / BIC	-399.30 / 808.6 / 827.8	-395.82 / 809.6 / 844.1	-391.20 / 808.4 / 858.2	-389.77 / 809.5 / 867.0
Observations (N)	341	341	341	341

Notes: Coefficients represent log-odds of ordered model.
Standard errors (SE) in parentheses. *p*-values from two-tailed tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.0$.
All VIFs for independent variables were below 2.5, indicating no problematic multicollinearity.

7.7.4 Mediation analysis: Testing model pathways

7.7.4.1. Mediation approach and proxies used

To assess whether digital and financial inclusion help explain the relationship between gender and reported challenges and impacts, I used the three-step mediation framework proposed by Baron and Kenny (1986) as a conceptual guide, testing (1) the total effect of gender on the outcome (e.g., needing help or experiencing loan challenges); (2) whether gender predicts the proposed mediator (e.g., limited financial inclusion); and (3) whether the inclusion of the mediator reduces the effect of gender on the outcome (Baron & Kenny, 1986). This method has limitations, including its reliance on the significance of the total effect and its assumptions regarding normality and sample size (Hayes, 2013). To address these limitations, I complemented this strategy with more robust estimation of indirect effects. Specifically, for cases where the gender effect was attenuated (either reduced or became non-significant) upon inclusion of the mediator, I used non-parametric bootstrapping (1,000 iterations in line with smaller samples) to estimate the indirect effect and associated 95% confidence intervals (Preacher & Hayes, 2004). This involved estimating: (a) the effect of gender on the mediator; (b) the effect of the mediator on the outcome (controlling for gender); and (c) the distribution of the product of coefficients across bootstrap samples. The proportion of bootstrap samples with an indirect effect in the opposite direction was used to approximate the p-value. This bootstrapped approach does not rely on normality assumptions and is well-suited for smaller or skewed samples. Sobel tests (Sobel, 1982) were also applied for reference, though these are known to be conservative in small samples. Finally, I checked for multicollinearity between gender and mediators; all VIFs were below 2.1, suggesting no inflation or redundancy in explanatory power.

My empirical measures of mediation focused on digital and financial inclusion. For digital financial inclusion, I used two proxies: (1) whether the Lender's app was the only digital loan app the respondent used, and (2) whether the respondent had taken only 0 to 1 loans through the app. These measures are admittedly imperfect (respondents may favor a single app or use informal sources), but they offer indicative evidence of digital credit exposure. Regression analysis confirmed that women were significantly more likely to rely exclusively on the Lender's app ($\beta = 0.76$, $p = 0.001$), while the number of loans taken was not significantly associated with gender. The proxy for digital inclusion (whether respondents used their own phone to access the app) lacked variation – as aforementioned, only 12 respondents reported using another's phone – and was not significantly gendered. This limits the variable's utility for mediation so it was not included.

Financial inclusion was captured using two proxies: (1) a composite variable I created for formal financial exclusion, which includes indicators of not having accessed formal banking, savings, or insurance that were drawn from the World Bank's definition of financial inclusion (World Bank, n.d.) ("formal financial exclusion"); and (2) a binary indicator of whether the respondent had other loan options ("broad financial inclusion"). While formal financial exclusion was significantly associated with gender, the latter broader financial inclusion variable was not, at least for women overall or rural women. It was, however, associated with female entrepreneurship. Therefore, broad financial inclusion was excluded as a mediator for challenge-related models but is considered in the impact-related models, particularly those focused on entrepreneurs.

While constrained by available indicators, this mediation framework enables a structured examination of whether women's elevated challenges in digital lending stem partly from systemic gaps in financial and digital inclusion, offering empirical and theoretical insight. These results should be interpreted as suggestive rather than causal. Future studies with experimental or longitudinal data could apply counterfactual-based mediation analysis to more rigorously estimate indirect effects.

7.7.4.2. Mediators in gender differences of challenges experienced

I hypothesized that women may experience greater challenges accessing and using the app, in part due to lower levels of digital inclusion, financial inclusion, digital literacy, and financial literacy. While I did find that women face greater challenges and more often need help, I now turn to examine if this is linked to lower levels of digital / financial inclusion or digital / financial literacy. To examine this, I focused on two key outcomes that were significantly influenced by gender: (a) needing more help, and (b) experiencing greater challenges downloading and using the app. I also explore one challenge where gender was significant: rural women experiencing not receiving a loan offer.

The first challenge I examine is whether help is needed. In a simple model, women were significantly more likely to say "no" to not needing help (i.e., they do need help) using the app ($\beta = -0.80$, $p = 0.042$). I then examined if women were significantly more likely to have the Lender's app as their only digital loan app, which they were ($\beta = 0.76$, $p = 0.001$). In the full model including both both gender and digital financial inclusion, the relationship between only having one app and needing help was statistically insignificant ($\beta = -0.50$, $p = 0.195$), while the gender coefficient weakened and was only marginally significant ($\beta = -0.71$, $p = 0.076$). This suggests that limited digital financial inclusion may partially mediate the relationship between gender and needing help with the app. A Sobel test of the indirect effect was not significant ($z = -1.213$, $p = 0.225$), suggesting that this variable does not significantly mediate the relationship between gender and needing help. Bootstrapping analysis (1,000 iterations) confirmed that the indirect effect was not statistically significant (95% CI includes 0; p = 0.214), providing no evidence of mediation.

In the mediation analysis for formal financial inclusion, I see women were significantly more likely to be financially excluded (β = 0.52, p = 0.028). When both gender and financial exclusion were included in the model, financial exclusion was a strong predictor of needing help (β = -1.18, p = 0.002), and the effect of gender weakened and became marginal (β = -0.67, p = 0.092). A Sobel test indicated a marginally significant indirect effect of gender through financial exclusion (z = -1.78, p = 0.075), providing suggestive evidence of partial mediation. Bootstrapping analysis (1,000 iterations) confirmed that the indirect effect was not statistically significant (95% CI includes 0; p = 0.318), providing no evidence of mediation. The evidence for mediation remains weak and inconclusive.

Turning to challenges downloading and using the app, I examined whether digital or formal financial exclusion mediated this outcome. Having the Lender's app as the sole digital loan app did not mediate the relationship. However, formal financial exclusion did. In the baseline model, women were marginally more likely to report difficulties downloading or using the app (β = 0.52, p = 0.062). When both gender and financial exclusion were included, the effect of financial exclusion remained strong and significant (β = 1.24, p < 0.001), while the gender effect became non-significant (β = 0.41, p = 0.16) (see Table 19). A bootstrapped mediation analysis (1,000 iterations) confirmed a significant indirect effect (Indirect effect = 0.652, 95% CI [0.039, 1.445], p = 0.036), providing strong support for mediation. This suggests that women's greater difficulty using the app is driven, in part, by their exclusion from the formal financial system.

Table 19. Mediation results: Gender and challenges downloading and using the app

Variable	**Model 1 (Total Effect)**	**Model 2 (Mediator)**	**Model 3 (Direct + Mediator)**
Female	0.52 (0.28)*	0.52 (0.24)**	0.41 (0.29)

Financial Exclusion			1.24 (0.29)***
Model Fit	Log-Likelihood: -218.36 AIC: 444.7; BIC: 460.1	Log-Likelihood: -209.29 Pseudo R^2: 0.011	Log-Likelihood: -209.07 AIC: 428.1; BIC: 447.3
Observations (N)	342	342	342

Notes: Coefficients represent log-odds. Standard errors (SE) in parentheses. *p*-values from two-tailed tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Models 1 and 3 are ordered logistic regressions and include threshold estimates. Model 1 thresholds: 1|2 = 1.75, 2|3 = 0.02, 3|4 = 0.43. Model 3 thresholds: 1|2 = 2.17, 2|3 = 0.06, 3|4 = 0.45. Model 2 uses a binary logistic regression; intercept = -1.06 (0.17)**.

Taken together, formal financial exclusion is a consistent and significant mediator of women's challenges, particularly in navigating and using the app. It is also marginally associated with women needing more help, whereas limited digital financial inclusion does not appear to play a meaningful mediating role in these outcomes.

Another notable finding relates to rural women, who were significantly more likely to report not receiving a loan offer compared to other women and men. A mediation analysis was conducted to test whether limited digital financial inclusion (measured as having only the lender’s app for digital loans) mediated the relationship between being a rural woman and experiencing challenges due to not getting a loan. While rural women were marginally more likely to have only one digital loan app ($\beta = 0.86$, $p = 0.060$), app access did not significantly predict loan challenges when included in the model ($\beta = -0.30$, $p = 0.523$), and the direct effect of being a rural woman remained strong and significant ($\beta = 1.72$, $p = 0.003$). A Sobel test confirmed that the indirect effect was not significant ($z = -0.64$, $p = 0.519$), suggesting that having one digital loan app does not mediate this relationship. A mediation analysis was conducted to examine whether financial exclusion mediated the relationship between being a rural woman and experiencing challenges due to not getting a loan. While rural women were significantly more likely to report such challenges ($\beta = 1.66$, $p = 0.003$), they were not more likely to be financially excluded ($\beta = 0.04$, $p = 0.931$), and financial exclusion did not significantly predict loan challenges when included in the full model (β = 0.68, p = 0.133). These results provide no evidence of mediation, as neither the association between the predictor and the mediator nor the indirect effect was significant.

Given the role of profit-maximization in algorithmic lending models discussed in the previous chapter, I tested whether indicators of borrower profitability and risk predicted loan offer exclusion. Neither having a history of late payments nor defaulting was significantly associated with being denied a loan offer. However, small loan size was: borrowers whose most recent loan was below the 25th percentile in size were significantly more likely to report not receiving a new loan offer (β = 1.04, p = 0.032). Importantly, the observed loan size refers to the borrower's most recent successfully obtained loan and may reflect being a newer borrower or proximity to underwriting thresholds.

A mediation analysis tested whether having a small loan size explains why rural women are more likely to report challenges getting loans. While both rural women (β = 1.63, p = 0.010) and small loans (β = 1.03, p = 0.038) were significantly associated with challenges in the full model, the association between rural womanhood and small loans was not significant (β = 0.27, p = 0.617). These results suggest that while both factors contribute to reported challenges, small loan size does not mediate the effect of rural gender identity. These findings indicate that while smaller loan sizes are independently linked to access challenges, they do not fully explain rural women's elevated odds of not receiving a loan offer.

7.7.4.2. Mediators in gender differences of impacts experienced

I also hypothesized that women may experience perceived greater benefits, in part due to lower levels of financial inclusion and access to resources. To examine this, I focus on three key outcomes that were significantly influenced by gender: (a) increased hopefulness about the future, (b) increased sense of control over money among female entrepreneurs, and (c) increased financial decision-making among female entrepreneurs.

Women overall

Focusing first on hopefulness, I conducted a series of exploratory mediation analyses to assess whether women's greater reported improvements might be explained by their patterns of (digital) financial inclusion. In the initial model, being female was marginally associated with higher reported hope (β = 0.40, p = 0.067). As aforementioned, having the lender's app as the only digital loan app used was significantly associated with higher hope scores (β = 0.50, p = 0.029). When both gender and sole digital loan app use were included in the model, the effect of gender weakened and lost statistical significance (β = 0.33, p = 0.13), suggesting a possible mediating role. A bootstrapped mediation test supported a positive but non-significant indirect effect (indirect effect = 0.113, 95% CI [–0.03, 0.28], p = 0.13). This suggests that women's greater hopefulness may be partially explained by their more limited use of other digital loan apps, though the bootstrapping results do not reach conventional thresholds of significance.

In contrast, formal financial exclusion appears to play a suppressing role. Women in the sample were significantly more likely to be formally financially excluded. When formal financial exclusion was added to the model with gender, exclusion was significantly associated with lower hope (β = -0.56, $p = 0.017$), while the gender coefficient increased and became statistically significant (β = 0.48, $p = 0.033$). This suggests that financial exclusion may be obscuring the full strength of the positive association between gender and hope. A non-parametric bootstrapping test confirmed this interpretation, revealing a significant negative indirect effect of gender on hope via financial exclusion (indirect effect = -0.141, 95% CI [-0.357, -0.004], $p = 0.04$). In sum women's greater financial exclusion may dampen the magnitude of their reported hopefulness, suggesting a suppression effect rather than mediation.

Taken together, these findings suggest that women's increased hopefulness may be partially mediated by more limited digital loan options (as indicated by sole app use), while formal financial exclusion appears to mask a stronger underlying relationship between gender and hope.

Female entrepreneurs

Female entrepreneurs were significantly more likely to report increases in hopefulness about the future and were also significantly less likely to report having other loan options compared to other women and men (β = -0.66, p = 0.03). However, when both variables (being a self-employed woman and having other loan options) were included in the same model, only the effect of being a self-employed woman remained significant (β = 0.96, p = 0.006), while having other loan options was not (β = 0.31, p = 0.17). This suggests that although self-employed women tend to have fewer alternative loan options, this does not mediate their increased sense of hopefulness. Having the lender's app as the sole digital loan app was not significantly associated with female entrepreneurs, nor was formal financial exclusion (hence not being assessed as a mediator).

Female entrepreneurs reported significantly greater improvements in their sense of control over money compared to other women and men. In the baseline model, being a self-employed woman was associated with higher odds of experiencing a positive impact on money control (β = 0.64, p = 0.035). When including access to other loan options in the model, self-employment remained a significant and positive predictor with slightly reduced magnitude (β = 0.60, p = 0.048), while having other loan options was negatively associated with financial control (β = -0.42, p = 0.05). To further assess potential mediation, I conducted a bootstrapped mediation analysis. The bootstrapped indirect effect was 0.27, with a 95% confidence interval of [-0.033, 0.711] and a p-value of 0.045, suggesting a modest but statistically significant indirect effect. This indicates that limited access to alternative loan options may partially explain why female entrepreneurs report greater improvements in financial control.

Female entrepreneurs reported greater improvements in financial decision-making. In the baseline model, being a self-employed woman was associated with higher odds of experiencing a positive impact on financial decision-making (β = 0.74, p = 0.022). When including having other loan options in the model, self-employment remained a significant and positive predictor (β = 0.72, p = 0.026). Having other loan options had a negative and insignificant association with financial control (β = -0.12, p = 0.57). Therefore, no significant relationship was found.

To ensure that the observed effects were not simply capturing something specific about being an entrepreneur more broadly or using the loan for productive purposes, I conducted a series of robustness checks. First, I examined whether self-employed men experienced similar benefits. Their coefficients were small and statistically insignificant across all three impact outcomes: decision-making ($\beta = 0.10$, $p = 0.719$), hope ($\beta = 0.04$, $p = 0.891$), and control over money ($\beta = 0.26$, $p = 0.346$) suggesting no meaningful effect. Although self-employed men were significantly more likely to report using the loan for work purposes, this did not translate into significant impacts in models that included both loan purpose and the outcome indicator.

Next, I assessed whether the observed impacts among self-employed women were mediated by how the loan was used (i.e., for work purposes). When loan purpose was included in the model, the effect of being a self-employed woman on perceived financial control declined from $\beta = 0.64$ ($p = 0.035$) to $\beta = 0.49$ ($p = 0.11$), and lost statistical significance, while the mediator itself was strongly associated with greater control ($\beta = 0.58$, $p = 0.008$), indicating partial mediation. A similar pattern emerged for decision-making: the initial significant effect ($\beta = 0.74$, $p = 0.022$) dropped to $\beta = 0.48$ ($p = 0.15$) when loan purpose was included, with a robust effect of the mediator ($\beta = 1.15$, $p < 0.001$). For hope, however, the effect of being a self-employed woman remained statistically significant even after including the mediator ($\beta = 0.72$, $p = 0.042$), alongside a strong effect of loan purpose ($\beta = 0.91$, $p < 0.001$).

These results suggest that among self-employed women, the loan's productive purpose partially mediates impacts on sense of financial control and financial decision-making, but not on hope. In contrast, self-employed men showed no significant impacts across outcomes, and mediation by loan purpose further attenuated their already weak associations. These findings underscore that the gendered effects of digital credit are especially pronounced among women entrepreneurs, some of which can be attributed to how the loan is used.

In sum, self-employed women appear to benefit more from digital loans in terms of hope, control over money, and decision-making. While having fewer alternative loan options is more common among these women, it only helps explain their increased sense of control, not their increased hopefulness or decision-making. Meanwhile, using the loan for productive purposes partially mediated the effect on financial-decision making and control, but not hope (see Table 20 summarizing mediation results among benefits). Self-employed men did not show significant differences in the impact categories.

Table 20. Summary of mediation results

	Increased hopefulness	**Increased sense of financial control**	**Increased financial decision-making**
Women overall	● **Only app use: partial** ● Formal financial exclusion: none *(suppression effect)*	N/A	N/A
Self-employed women	● Alt. loan: none ● Loan for work: none	● **Alt. loan: partial** ● **Loan for work: partial**	● Alt. loan: none ● **Loan for work: partial**

7.8. Discussion

This chapter explored the benefits and challenges users experience in accessing and using ML-based alternative lending tools, and how these compare between women and men. I employed a multi-method analytical strategy grounded in my conceptual framework to identify findings drawn from survey data and complemented findings with analysis of a large Google Play review dataset. The results speak to my guiding hypotheses.

- H1 predicted that women, especially rural women, would experience greater challenges in using the app. This hypothesis is supported, though the strength of evidence varies by challenge type: women report a greater need for support and rural women are more likely to report not receiving a loan offer, while reported difficulty using the app is only marginally higher for women. For other specific challenges, no statistically significant gender differences were observed. H1b posited that challenges accessing and using the app would be linked to lower levels of digital and financial inclusion. This is partially supported: only formal financial exclusion mediated the relationship between gender and reported challenges.

- H2 predicted that women, particularly self-employed women, would report greater relative benefits from using the app. This is partially supported: one specific benefit (hopefulness) was marginally higher for women compared to men, but most benefits did not differ by gender. Self-employed women reported greater perceived benefits across a range of outcomes. H2b expected that gender differences in benefits would stem from women's more limited access to financial alternatives. This is partially supported: while some of the reported benefits appear to be linked to women's more limited access to financial alternatives (e.g., lack of other loan options among female entrepreneurs), other findings suggest a more complex pattern. For example, while formal financial exclusion is more common among women, it suppresses rather than explains their higher levels of hopefulness, indicating that women report greater hope despite their exclusion from formal finance.

In particular, four key findings emerge that illuminate gendered dynamics from my hypothesis-driven analysis: (1) Women need more help and experience greater challenges than men in using the app; (2) Rural women are significantly less likely to receive loan offers, a pattern not fully explained by digital or financial exclusion, but associated with smaller loan sizes; (3) Female entrepreneurs report greater perceived benefits using the app and accessing loans, partially mediated by lack of alternative loan options; and (4) ML-based lending tools expand access to finance, yet women appear included on unequal terms. While the findings reflect observed associations rather than causal effects, they offer important insight into the ways that benefits and challenges of ML-based credit assessment tools are gendered. Each of these findings is discussed in detail below, in relation to my hypotheses and conceptual framework.

7.8.1. Fairly easy to use, though greater challenges experienced by women

Overall, most users reported that downloading and using the app was relatively easy. However, women were more likely than men to report challenges downloading and using the app. In chapter 5, I outlined fintech perceptions for gender differences in loans, which included women facing greater challenges in using the app. This perception is confirmed by the data in this chapter. In a model controlling for demographics, digital loan app usage, and financial inclusion, the gender effect on experiencing challenges weakened and became non-significant, while financial inclusion was found to be significant. This suggests that gender alone does not fully explain these difficulties once broader structural and socioeconomic factors are considered.

To further investigate and test whether challenges are associated with relatively lower digital and financial inclusion and literacy (H1b), I examined the mediating role of formal financial exclusion. Results showed that women were significantly more likely to be financially excluded, and financial exclusion was a strong and significant predictor of reporting difficulties using the app. When both gender and financial exclusion were included in the model, the gender effect became non-significant, while financial exclusion remained significant. A bootstrapped mediation analysis confirmed a significant indirect effect providing strong evidence that formal financial exclusion mediates the relationship between gender and app-related challenges. This finding is consistent with existing research showing that individuals who access formal financial services tend to have heightened financial literacy and behaviors, versus those who are financially excluded. Indeed, access to financial services not only facilitates financial transactions but can improve financial literacy and behavior (Herrerias & M. O. Alvarez, 2023). Increased exposure to financial activities is similarly associated with improved financial knowledge, with links particularly pronounced among women (Reddy et al., 2024).

Relatedly, there were gender differences in needing help. With an odds ratio of 0.45, the odds of women *not* needing help were 55% lower than those of men, indicating that women were significantly more likely to need help when using the app. These trends of women needing more help in using digital financial tools track with existing data around women needinrg help using mobile money. The World Bank's Findex Database finds that only 58% of women can use mobile money without help, compared to 84% of men (World Bank, 2021). Once key demographic and financial inclusion factors were included in my model, the gender effect on needing help lost significance. Mediation analysis initially suggested that financial exclusion may play a partial mediating role in this outcome as well, though further analysis revealed this to be limited.

These findings support my hypothesis that women face greater challenges in accessing and using the app. Women were significantly more likely than men to report difficulties downloading or using the app and to indicate that they needed help navigating it. These disparities appear to be partially explained by differential access to formal financial systems: once financial inclusion was accounted for, the gender effect was attenuated. In models including only demographic controls, higher education was a significant predictor of reduced need for help and fewer usage challenges, suggesting that education plays a moderating role. Other studies have found that higher levels of education are associated with increased probability to use digital credit for women in Kenya (Johnen & Mußhoff, 2023). However, the effect of education lost significance in fuller models that incorporated digital loan app usage and financial inclusion variables, indicating that digital loan app usage and financial inclusion may better explain these outcomes than education alone.

A technofeminist perspective pushes my analysis further by questioning how the app itself may be designed with gendered assumptions about the default or "ideal" user. Drawing on technofeminist theory (Wajcman, 2006), fintech apps may embed implicit norms about an average user (e.g., their digital fluency, familiarity with financial products) that align more closely with men's experiences, particularly those in formal employment or urban contexts. The app's interface, language, and onboarding processes may not adequately account for women's more varied patterns of access, including relatively lower digital and financial access and literacy. In this sense, usability challenges are not simply user-side limitations, but also reflect design choices that fail to accommodate diverse users and gendered realities.

Fintechs can take various approaches to reduce challenges women face and support more equitable access. First, fintechs may explore design tweaks to the app to enhance its ease of use, including doing user experience (UX) design updates particularly informed with and for women, including women in rural areas. Doing so can result in fintech products that work better for women (International Finance Corporation, 2024). The apps may further benefit from greater financial literacy integration. Financial literacy significantly amplifies access to and benefits of digital financial services to support women's economic empowerment (Showkat et al., 2025). In short, results suggest that using a gender lens in the design of the technology and providing financial literacy support throughout the loan lifecycle can mitigate issues and enable fintechs to better support underbanked populations.

7.8.2. Rural women and differing receival of loan offers

In examining gender differences in specific challenges related to the app, loan access, and repayment, only one specific challenge was found to be significant: not receiving a loan offer among rural women. Rural women have 5.24 times the odds of not receiving a loan offer compared to others. This pattern aligns with the descriptive statistics: While 6.4% of the full survey sample reported not receiving a loan offer (7.5% of women, 5.5% of men), this rose to 22.72% of rural women ($\chi^2 = 8.49$, $p = 0.0143$).

This effect became even more significant in the third model incorporating demographics, digital loan app usage, and formal financial inclusion. However, the interaction term became insignificant in a fourth model once the main effects of gender and rural residence were included, indicating that the increased likelihood among rural women reflects the combined, additive effects of gender and rural residence rather than an additional interaction effect. In running two different models for men and women, rural residence was significant for women but not for men, in regards to not receiving a loan offer.

In examining potential mediators, neither limited digital financial inclusion and formal financial exclusion variables explain the disparity or mediate the relationship between rural women and experiencing this challenge. Default and late payment behavior also do not explain these differences. The lack of other variables being able to explain this difference highlights potential issues in the algorithms and on the supplier side.

Prompted by my findings from Chapter 6 which discuss the role of profit-maximization in algorithmic lending models, I tested whether indicators of borrower profitability and risk predicted loan offer exclusion. Neither late payments nor default status were significant, but small loan size was. While both rural womanhood ($\beta = 1.63$, $p = 0.010$) and small loan size ($\beta = 1.03$, $p = 0.038$) were associated with not receiving a loan offer, small loan size did not mediate the effect of being a rural woman.

Taken together, these findings suggest that the observed disparity is not explained by measured user-side characteristics including lower financial inclusion and literacy, and instead raise questions about supplier-side processes and potential issues in the underlying algorithms.

7.8.3. Perceived benefits and heightened impacts for female entrepreneurs

Overall, respondents reported positive impacts across a range of perceived empowerment and financial well-being metrics. These findings echo other research that finds modestly positive effects of subjective well-being for borrowers from ML-based credit assessment apps (Björkegren & Grissen, 2020), as well as a study in Kenya where digital credit facilitated through an ML-based credit assessment app improved borrowers' financial well-being, with greater impacts among borrowers with limited access to credit and those who take loans for business purposes (A. Y. Chen et al., 2025).

I hypothesized that women, particularly women entrepreneurs, may benefit more than men tied to women's relatively more limited access to finance. Findings reveal that benefits are largely similar for men and women who access the apps, though subgroup analyses revealed that female entrepreneurs reported significantly greater impacts in three areas: hopefulness about the future, financial decision-making, and sense of financial control.

Regarding hopefulness about the future and improved financial decision-making, the effect for female entrepreneurs remained significant when controlling for demographics, digital loan app usage, and financial inclusion. However, when gender and employment status were added in a fourth model, the effect of self-employment among women dropped to marginal significance. Other indicators emerged as significant predictors, including exclusive use of the Lender's app and formal financial inclusion. For the impact on one's sense of financial control, the third model incorporating financial and digital inclusion variables, resulted in a drop to moderate significance. The exclusive use of the Lender's app was significant.

Mediation analysis showed that both loan purpose and lack of alternative loan options partially mediated these relationships for self-employed women. Specifically, lack of alternative loan options partially mediated increased sense of financial control, but not on financial decision-making or hope. Using the loan for work purposes explained some of the effects on sense of financial control and financial decision-making, but not on hope.

In short, female entrepreneurs tend to experience greater perceived benefits, which is partially mediated by lack of alternative loan options. Several impacts are also mediated by using the loans for work purposes. Hope remained unexplained by the mediators tested.

My findings suggest that the intersection of gender and entrepreneurship is an important lens for understanding who benefits most from digital credit. While Chen et al (2025) found that the loan purpose mattered for the experience of impacts, they did not find overall gender differences. Importantly, gender difference only emerged for me in subgroup analyses (e.g., among entrepreneurs), illustrating the importance of intersectionality research. Future research could benefit from exploring gender-differentiated impacts of digital credit on entrepreneurs in particular.

These results should be taken with caution, given both the various survey limitations (discussed in 7.9) and learnings from the microfinance industry. Despite its theoretical promise, studies on microfinance have been mixed regarding whether women experience greater benefits than men, although microcredit has long been thought to be particularly valuable for women who represent the majority of microfinance borrowers. This was hypothesized linked to the larger financial inequities that women (particularly women entrepreneurs) face and that access to credit would shift the balance of power in households, while supporting women through enhanced autonomy (Horton, 2017). Studies have been mixed on their evidence for providing greater benefits to women, with benefits of microfinance being partially mitigated due to their for-profit approach and tensions with the wider mission of women's empowerment (Horton, 2017).

7.8.4. Financial inclusion increases, yet gender nuances persist

Exploratory findings in this chapter offer insights connected to the other chapters. First, interviewees discussed that ML-based credit assessment is effective in providing financial access to people who are unbanked and underbanked. The exploratory findings in this chapter highlight that many of those coming to the apps are outside of formal financial folds: 31% of the sample report not having access to any formal financial services including loans, savings account or insurance (25.% of men and 36.8% of women). The ability for these ML-based credit assessments to enhance access to credit to people outside of formal financial folds has been found in prior research (Björkegren & Grissen, 2018; Chioda et al., 2024).

While the app is clearly reaching those outside formal financial systems, users in this sample are more financially included than the national average. This raises questions about who these tools are reaching and who may still be left out. Importantly, limited access to formal financial services is associated with greater difficulty using the app and helps explain the greater challenges reported by women. This suggests that fintechs seeking to promote inclusion should go beyond access and support financial literacy within the app or as part of the design.

Second, both interviews and survey data indicate that women tend to receive smaller loans than men, while women have better repayment behavior. Importantly, in the survey, while more women noted fewer late payments and default, gender differences were not statistically significant. Gender differences in loan size are significant, and it remains unclear whether women are offered smaller loan sizes or are more likely to opt for them when given a choice. Still, this pattern mirrors findings in the microfinance literature, which consistently shows that women borrowers tend to be more reliable despite receiving lower amounts of credit (D'Espallier et al., 2011). Descriptive statistics show that women have slightly higher rates of not receiving a loan offer, which are significantly higher for rural women as aforementioned.

Taken together, these findings suggest that while ML-based credit assessment tools expand access, they do not effectively address underlying gender gaps in financial inclusion. This aligns with prior research finding that formal digital credit, claiming to be gender neutral, led to an increase in the gender gap in financial inclusion in Kenya (Johnen & Mußhoff, 2023).

7.9. Conclusion

This chapter provides important insights into benefits and challenges experienced by users of ML-based credit assessment apps through a case study of one fintech company in Kenya. The findings show that while these tools expand access to credit for those excluded from formal financial systems, they do not close gender gaps. Women face greater barriers in navigating the app and tend to receive smaller loan amounts, while rural women receive fewer loan offers. Meanwhile, female entrepreneurs experience heightened impacts overall. While fintech innovations hold promise, they fall short in addressing gender inequitable access to finance.

My findings affirm insights from feminist economics, which argue that access to resources, even limited, can produce meaningful empowerment for structurally disadvantaged groups. They also highlight a concern elucidated through technofeminist critique: "neutral" approaches in tech design can inadvertently privilege male-coded financial behaviors and background, risking the reproduction of inequalities they claim to disrupt. Taken together, these results offer partial support for my guiding hypotheses: gendered inequalities persist in both access to and outcomes from digital credit tools. At the same time, the analyses reveal that some expected explanatory mechanisms (i.e. as women's financial exclusion) did not consistently account for observed differences, underscoring the complexity of these dynamics. Overall, these patterns suggest that financial technology is not inherently equitable, but must be deliberately shaped to reflect and take into account diverse social and economic realities.

Some may argue that fintechs are not obligated to make their platforms gender-sensitive; after all, these tools are reaching some users and expanding access more broadly than traditional banks. From this perspective, if women happen to face greater challenges or receive smaller loans due to pre-existing financial exclusion, that is unfortunate but not necessarily the fintech's responsibility (and addressing these issues head on could be interpreted as an act of algorithmic affirmative action). This view overlooks two points. First, the absence of gender-sensitive design is not neutral. If credit algorithms and user interfaces are built around behaviors more typical of men (e.g., larger loans, higher financial literacy) then the system inadvertently privileges those profiles. This doesn't merely reflect inequality; but can deepen it. Second, gender-sensitive design is not about favoring one group, but rather expanding usability and effectiveness across a range of users. The goal is not to make gender exceptions, but to recognize that current systems are already optimized around a narrow user profile. By incorporating a gender lens and recognizing the paradox of "neutral" design approaches, fintechs have the opportunity to further their social missions. On the other hand, not incorporating gender-sensitive design can result in social setbacks by potentially worsening gender gaps in financial inclusion. Critical questions remain around how profit priorities are limited in their power to transform societal inequalities.

The findings in this chapter elucidate recommendations for the Lender and other fintechs. First, the Lender should conduct gender-disaggregated UX design exercises to understand how to make the app more approachable and usable for women, as well as for people broadly who are newer to digital financial apps and financial inclusion. Second, they can provide simple, financial literacy educational tips on the app informed by a gender lens. An additional recommendation is to the conduct a gender analysis of the company's data in regards to who is receiving loans and at what amounts, while examining factors in the model that are most predictive in assigning credit scores, to examine if there is an unintentional bias occurring that results in women getting smaller loans and rural women being denied loans more frequently holding other variables constant.

There remain various limitations. Importantly, I coded gender as binary (no respondents indicated being nonbinary despite this being an option in the survey), but I recognize that gender is a spectrum and acknowledge this limitation. Future research should further explore impacts for nonbinary people. Second, the sample size for the survey is on the smaller end of 342 responses, limiting generalizability and reducing statistical power, particularly for subgroup analyses. Also, there are notable differences between male and female respondents in the sample (e.g., age, marriage status), as well as between the overall sample and national statistics for Kenya (e.g., education level, financial inclusion status). While the sample may be considered reflective of a subset of app users, it is not generalizable to all potential or current users of digital credit apps in Kenya. Additional research can further examine, through a gender and intersectional lens, which populations are accessing these apps and which groups may be underrepresented or excluded and why. Relatedly, respondents self-selected into the survey, introducing potential response bias. Although this was partially mitigated through Qualtrics controls and targeted outreach by the Lender to ensure demographic diversity for key variables I identified, it remains an important limitation to acknowledge. Also, participants self-reported their rural, urban or peri-urban status which may reflect their own perceptions.

In regards to the modeling approach, while the analysis was guided by a theory-driven and sequential modeling strategy, the decision to proceed with additional models only when baseline associations between gender and the outcome were statistically significant may obscure effects that emerge only after adjusting for covariates. This reflects an intentional emphasis on identifying total (unadjusted) gender differences, but may limit detection of suppressed or conditional effects. Additionally, while the estimation of multiple models introduces a risk of false positives due to multiple comparisons, the models and variables were theoretically motivated and the analytical approach was grounded in my conceptual framework. This mitigates the risk of p-hacking, and results were interpreted with attention to effect sizes and consistency across model specifications. Finally, the analysis is based on cross-sectional, self-reported data, which limits the ability to make causal inferences. Unobserved variables (e.g., risk tolerance, confidence) may influence app usage and perceived impacts, raising the possibility of endogeneity or selection bias. Therefore, mediation analysis findings should not be interpreted as demonstrating causal pathways, rather they provide insight into mediating mechanisms consistent with my theoretical expectations. These findings offer a foundation for future work, including with larger samples. Future research may include experimental designs to assess mediation and can further explore relationships outlined in this chapter, including intersectionality related to how women experience challenges and benefits of ML-based credit assessment apps.

CHAPTER 8: CONCLUSION

8.1 Discussion of Findings

This thesis explored whether ML-based credit assessment tools by fintechs reinforce or mitigate gender inequitable access to finance in LMICs. I had two sub-questions: (1) In what ways do the underlying logics, design choices, and management decisions of ML-based credit assessment tools by fintechs embed or challenge gender biases, and how do these choices influence gender equity in access to finance? (2) What benefits and challenges do users experience in accessing and using ML-based credit assessment tools, and how do these compare between women and men? Chapter 5 and 6 addressed RQ1, whereas Chapter 7 addressed RQ2.

In Chapter 5 I find that underlying logics, design choices, and management decisions of ML-based credit assessment tools by fintechs take a "neutral" or gender "blind" approach. These gender "blind" approaches are grounded in beliefs that ML is objective and data reflects the truth. This leads to a lack of grappling with the ways data, features for creditworthiness, and access to apps are gendered. Overall, tools appear to increase access to finance, but not gender equitably: Interviewees consistently report that less women access loans and receive lower amounts than men, despite being better repayers. Fintechs identify demand- and supply-side reasons for gender differences, but frame them as outside their responsibility. However, that women are observed as better repayers reveals a market inefficiency and potential discriminatory effect, further linked to profit optimization objectives. This chapter introduces the concept of *encoded gender norms*, whereby without explicit attention to the gendered nature of data and algorithmic design, AI tools reproduce existing inequalities. In doing so, they reinforce gender norms as self-fulfilling prophecies. The idea that AI is inherently objective and, when left alone, fair, is seductive and misleading. In reality, algorithms reflect the perspectives, priorities, and values of the people and institutions that design them.

In Chapter 6, I illustrated how fintechs' varying and often limited definitions of fairness – whether framed in terms of technical accuracy, social impact, or accessibility – inform the design and management of ML-based credit assessment tools, and fail to challenge gender inequities in access to finance. I identified perceptions of fairness in fintechs from both a process ("fair" algorithmic design) and outcome ("fair" credit assessment) perspective. From a process perspective, the prevalent notion of "fairness through unawareness" assumes that removing demographic data eliminates bias. However, this ignores structural inequities embedded in alternative data sources (e.g., mobile usage, digital footprints) and features (e.g., income), as well as how these may act as proxies for race, gender, and class. From an outcome perspective, fintech actors often define fairness as achieving social impact or being benefit-maximizing, prioritizing inclusion over equity. Others equate fairness with accuracy, reducing ethical concerns to mathematical optimization. These perspectives align with Silicon Valley logics that emphasize scalability and profit over social justice. Consequently, algorithmic lending reinforces existing gendered financial disparities while obscuring normative choices. Crucially, fairness in ML-driven lending is not dictated simply by individual ethics, but rather institutional priorities and economic logics. I introduce the concept of *situated priorities* which posits that priorities are constructed based on specific contexts, pressures, and epistemological standpoints, leading to varied fairness and responsibility practices. This chapter informs the persistence of gender inequitable access to finance in ML-based alternative lending, but also raises important questions about how incentive structures shape fairness approaches: fintechs are not acting irrationally within their market and regulatory environments.

Taken together, underlying logics – including belief in data as the truth and ML as objective – combined with epistemological positions and organizational priorities influence design choices and management decisions in ways that are vulnerable to embed, rather than challenge, existing gender biases. This occurs not through intentional exclusion or algorithmic performance disparities necessarily, but through a failure to recognize or challenge the ways algorithms intersect with existing gender norms and inequities. The technologies become prone to *default discrimination* whereby neutral approaches replicate discriminatory patterns that exist in society. Interviewees report that more men than women access loans and at higher amounts that they identify are linked to several demand- and supply-side considerations, which are shaped by gender norms and structural inequalities. I present the concept of encoded gender norms to extend the concept of default discrimination more specifically to the context of gender, recognizing that technology, data, and features in models are gendered and taking a neutral approach results in reflecting and encoding existing gender norms and structural inequalities.

While algorithmic tools may reduce some forms of overt bias (e.g., discrimination by loan officers or through collateral requirements), they do not necessarily address – and can even obscure – the deeper, insidious gendered exclusions embedded in financial systems. In doing so, the technologies risk normalizing unequal outcomes under the guise of objective machines. That women are considered better repayers but not getting similar loans illustrates that the promise of meritocracy pursued under American banking logics is not achieved. How fairness is perceived rationalizes gender differences in algorithmic facilitated lending, including through priorities of scale and profit. Meanwhile, normative choices and tradeoffs between inclusion and equity are obscured. These value-laden decisions shape how and for whom access is expanded, without confronting the gendered structures that underpin financial exclusion.

In Chapter 7, I explore RQ2 to unpack gender differences in benefits and challenges experienced by users when accessing and using ML-based credit assessment tools through a case study of one fintech in Kenya. While these tools expand access to credit for those excluded from formal financial systems, women face greater challenges in using the app, need more help, and tend to receive smaller loan amounts. In addition, rural women receive fewer loan offers. Meanwhile, female entrepreneurs experience heightened impacts, which is partially mediated by a lack of alternative loan options. These patterns of benefits and challenges reflect both the value these technologies are providing (including to incredibly vulnerable people), and reinforce a finding in chapter 5 whereby fintechs discuss one reason for gender differences in lending being that women are facing more challenges and needing more help using the apps. Given the various limitations outlined in Chapter 7, results may best be understood as exploratory, rather than representative. Still, they remain valuable as they offer important insights into a relatively underexplored area and represent an advancement in existing research.

I return to my overarching question: Do ML-based credit assessment tools by fintechs reinforce or mitigate gender inequitable access to finance in LMICs? While fintech innovations hold promise and are supporting access to finance overall, they fall short in addressing gender inequitable access to finance. Developers and managers of these technologies do not consider or adequately address how men and women access technology differently and the ways gender inequality can be reinforced in algorithms and how they are managed. Furthermore, organizational priorities and market incentives influence technology design and management with implications towards more profitable (gendered) segments. AI technology is not inherently equitable, and "neutral" approaches inadvertently privilege male-coded financial and digital behaviors and backgrounds. Despite positive perceived benefits experienced by people accessing the tools and the enhanced access to finance the technology is facilitating, gender differences persist in that access. Given the ways that values are embedded, my research illustrates that ML-based alternative lending tools could, however, be deliberately shaped to reflect and take into account diverse social and economic realities. Algorithms are not necessarily good or bad; rather, they are tools that can be built and deployed in different ways.

8.2 Interpretations

While this thesis does not attempt to quantify the degree to which ML-based credit systems reinforce gendered economic inequities, it instead focuses on the "ground truth" regarding the underlying logics of developers and managers who design and oversee these tools. This allows for an examination of how concepts of inclusion and equity are understood, addressed, or sidelined in the processes of algorithmic design and management. The thesis also incorporates user voices to understand their own perceived challenges in and benefits from accessing the technology. The following interpretation explores the implications of these framings and their findings.

8.2.1. Beyond bias to reproduction of structural inequities

The concept of bias in ML systems is often misunderstood or inconsistently applied. This is partly due to there not being alignment on its definition, similar to debates around fairness. Much of the technical literature defines bias narrowly in terms of statistical disparities (e.g., differences in error rates or performance metrics) across demographic groups. This is typically tied to underlying training data and the over- or under-representation of different identities therein (Buolamwini & Gebru, 2018). While this is very important to understand and unpack, the broader academic field has a more expansive view of bias in AI that includes reinforcement of social norms, systemic discrimination, and broader patterns of inequity (*Scientific Consensus on AI Bias*, 2025).

In this thesis, I intentionally adopted an open and sociotechnical perspective regarding if and how bias may manifest, including a focus on how ML-based credit assessment tools may reflect or amplify gender inequities in financial systems. This orientation was shaped in part by information I was able to access, the interview data that emerged (where participants explained gender differences in lending outcomes by pointing to the proxies and features used in models), and user-side constraints. Interviewees highlighted optimization goals and business priorities as key drivers shaping tool design and management, offering important insights into how fairness is interpreted and implemented in practice. As a result, I do not assess bias related to performance and accuracy metrics. Instead, and in line with my theoretical framework, I examine a different but important form of bias: how algorithmic systems encode, reflect, and potentially reinforce gender norms and inequities. This approach recognizes that concepts such as "fairness" and "accuracy" are not neutral, but normative. Beyond requiring access to underlying data on the model and borrowers, attempting to unpack performance discrepancies in ML-based credit assessment would be a normative exercise, one that requires decisions about what constitutes fair or accurate outcomes, which I purposefully do not attempt to define or evaluate as much as observe.

In the case of credit assessment tools, gender bias does not necessarily arise from overt or intentional discrimination; rather the use of variables reflects broader social and economic inequalities that can become embedded in the tools. Features such as income, employment history, or spending patterns are used to assess creditworthiness – yet these same features are shaped by systemic disparities that disproportionately affect women and other marginalized groups. In this way, inequity is reproduced not intentionally, but through the logic of data-driven decision-making itself. This echoes research in other domains, whereby proxies connected to marginalized groups are used in algorithms resulting in discrimination (Obermeyer et al., 2019).

How fairness is defined and operationalized becomes critical to unpack how fintech actors justify and sustain differences in outcomes. Decisions around which fairness metrics to use (and how to weigh them against priorities like accuracy or profit) are deeply normative, even when presented as technical. When fairness becomes subject to business goals, systems risk reinforcing inequality while appearing objective and efficient. This meets the needs of the business, but not necessarily societal goals. This highlights that a bigger picture view of "bias" in AI requires examining structural inequalities and interrogating the underlying personal and institutional logics that guide system design and management.

While based on self-reported accounts, the perceptions of and reasoning for gender differences expressed by interviewees are analytically significant. They provide critical insights into how gender disparities are observed, understood, and rationalized within fintech organizations. These perceptions reveal not only the presence of gendered differences in lending patterns, but also that fintechs recognize that gender norms and inequities creep into algorithms. This qualitative data thereby helps surface the internal understandings and logics, while also revealing interviewees' nuanced understanding of the ways that gender inequities are actually embedded in the technology itself. It reflects how fintechs don't necessarily see the reproduction of structural inequities as harmful "bias" that must be addressed in and through the tools. It opens up reflexive moments around questions such as whose responsibility it is (or not) for addressing inequity and social justice.

8.2.2. Replacing one form of exclusion with "rational" others and encoding gender norms

With persistent and gendered divides in financial inclusion, the advancement and proliferation of ML-based credit assessment tools is often framed as a mechanism to overcome barriers to access and unlock progress. Indeed, these systems can address certain historical exclusions – such as gender discrimination by loan officers. However, my findings suggest that while one form of exclusion may be mitigated, others are introduced in its place. These include exclusions rooted in digital divides, opaque decision-making, gendered proxies, and embedding of social inequities into seemingly neutral algorithmic tools. The result is that gendered behaviors and disparities – such as differences in access to employment and income levels – may be codified and output in credit scores that are seen as both "accurate" and rational. Embedding social norms into machines makes them appear more objective, while also harder to detect and challenge. As a result, ML systems risk legitimizing and perpetuating gender inequalities as rational. This aligns with prior research that digital credit increases gender gaps in financial inclusion, which the researchers found was linked to differences in socio-economic variables (Johnen & Mußhoff, 2023).

ML-based credit assessment tools can therefore play a role in shaping gender, by encoding gender norms and inequalities as rational. Unlike Wajcman's "cyborg figure" that offered women freedom from biological sex differences (Wajcman, 2006), algorithmic technologies encode gendered behaviors. These could have unintended consequences for both women and men: women are deemed less creditworthy, while men are offered riskier loans and potentially more so exposed to cycles of debt. While gender and gender roles are malleable and change over time and space, algorithmic technologies can reinforce certain notions of gender under "objective" technology outputs. ML tools present outputs as truth, legitimizing their power. These technologies can thereby present gender norms as truths. This narrative can be reinforced over time under systems that reward certain behaviors.

8.2.3. The persistent "black box" problem

Even if such systems represent an improvement over traditional lending practices, their opaque or "black box" nature risks solidifying existing inequalities under a veneer of technological objectivity. Developers and managers may view the tools as more fair or less biased than human decision-makers, but without transparency or meaningful accountability, such tools can normalize and automate discriminatory outcomes, especially if there is a limited sense of responsibility in addressing the ways that gender is encoded into data and algorithmic systems. This raises a critical normative question: Is better than the status quo good enough, especially if the tools are inscrutable to those they affect? The findings suggest that without deliberate attention to gender equity, even marginal gains may come at the cost of reinforcing structural disadvantages in insidious ways over time. Moreover, perceptions that the tools are "neutral" may reduce attention and resources to addressing underlying structural inequalities.

8.2.4. Normative and institutional constructions of fairness

This prompts deeper reflection on what counts as fair in algorithmic systems and, critically, who should get to decide. The default is that developers and managers decide what is deemed fair. However, these decisions have consequential and often unexamined impacts on vulnerable communities in the "AI for Good" space. Delegating these normative decisions to fintech actors without public scrutiny or engagement risks embedding definitions of fairness that are captured by corporate logics and certain situated priorities. Silicon Valley logics of machine learning as objective masks these powerful normative decisions, and associated ideas about what metrics matter, what tradeoffs are acceptable, and what – and whose – outcomes are prioritized. When occurring in opaque ways, even marginal improvements can still reproduce or entrench social inequities under the guise of being "for good". The question is not necessarily, what does fairness mean and look like, but rather, who gets to define fairness and ensure systems are accountable? In *Voices in the Code,* Robinson highlights both the opportunities and challenges of making such normative decisions more democratic and accountable (D. Robinson, 2022).

These insights can be applied to AI applications in international development and those “for good” and are further discussed in the future research agenda section.

These squishy grounds illustrate that technologies are not objective, but deeply value-laden tools. Even the decision to remain “neutral” or gender “blind” reflects particular values and priorities. Perceived neutrality is itself a moral positioning. Fairness cannot be solved for through technical strategies, rather it requires making choices. These choices shape who is deemed creditworthy under what – and whose – priorities and conditions.

8.2.5. Not a deliberate wrongdoing by fintechs, but a critical role for policymakers

My findings illustrate that there is a critical role for policy and public oversight to ensure that questions of equity are surfaced and addressed not simply by private actors, who have their own needs and priorities. I am not claiming that there is any deliberate wrongdoing by private actors. On the contrary, developers and managers are operating from their own positionality and situated priorities that are reinforced by investors and funders. Various interviewees discussed how they are innovating to be more inclusive and approaching sticky ethical questions, while also expressing hesitations, reflexivity, and – in some cases – calls for help, in grappling with this emerging technology where policy guidance and standards are lacking. Also, fintechs are, for the most part, fulfilling their organizational goals by optimizing for efficiency, profit, and market expansion within existing legal and regulatory frameworks. This is to be expected by for-profit institutions. A deeper issue lies in the absence of strong policy guidance and accountability structures. Without clear standards for fairness, equity, and transparency, private actors are left to define these terms on their own, often in ways that – unsurprisingly – align with business interests rather than social justice. They are, after all, businesses not social justice warriors.

My thesis underscores the critical role of policymakers and regulators in shaping the ethical and structural boundaries within which AI systems are developed and deployed. When left unchecked, even well-intentioned innovations can reproduce or deepen social inequities – not because of malice, but linked to situated priorities, a lack of standards and awareness, as well as a systemic gap in governance. This is a particularly important finding and interpretation as we are in a current time of global hesitation and pushback on regulating AI, with general sentiment of leaving the technology unregulated in order for countries to capitalize on its benefits and innovations potential. My findings illustrate that as AI is a technology proliferating via for-profit institutions, regulation is critical to provide guardrails for normative questions that algorithmic tools inevitably face and negotiation of priorities for the public interest.

The development industry also plays a role by prioritizing scale and focusing on investing in for-profit ventures. There's less interest in supporting innovations that prioritize depth (versus scale) and process (versus output). Development funders are important stakeholders that, similar to investors, influence priorities.

8.2.6. Perceived benefits and limitations

This thesis shows that ML-based credit assessment tools can and do lead to perceived benefits for many people, particularly pronounced among self-employed women who have limited options for financial access. Furthermore, the tools are helpful and in some cases, critical, for emergencies. Still, risks persist such as exposing borrowers to over-indebtedness or dependency on the app. Furthermore, they do not address the deeper economic and financial issues that can exist in LMICs. As such, this is a story not simply of success or failure. Rather, it's a story of nuance: these tools represent a form of progress, but one that is incremental, uneven, and shaped by existing power structures. They offer real benefits to people and enhance access to finance to people that need it, but don't fundamentally alter the economic conditions that produce vulnerability in the first place, nor do they address persistent structural gender economic inequities.

8.2.7. Inclusive normative decision-making and potential tensions

Although this did not emerge directly from interviews, I wonder: Could it be that the algorithm disbursing fewer or smaller loans to women is (in some contexts) acceptable to women and other end users, as long as others in their household or community still gain access? Some women may view financial inclusion through relational or collective lenses, finding value in shared outcomes, even if they themselves receive less. This possibility underscores the importance of understanding fairness from the perspective of users, particularly those who are more vulnerable or marginalized. What is seen as fair, empowering, or beneficial may differ across social and cultural settings, and may not always align with individualistic models of empowerment or success.

This raises a further question: Even if such patterns are socially accepted, should they be? Does their normalization reinforce gendered assumptions about who is creditworthy or economically deserving? This tension points to the need for critical engagement with how fairness is defined in different contexts – not only by system designers, but also by the communities and societies in which they live and operate.

It also highlights a larger question around the social acceptability of using algorithms to score people. Is this an appropriate application? Is the relative benefit worth the potential consequences? As accuracy in credit scoring can be a self-fulfilling prophecy, assessing bias and accuracy remains a daunting and value-laden task. Even if incorporating community voices, without changing the higher-level priorities and incentives, risks can persist. This brings to light a 2025 evaluation by MIT Tech Review journalists on an Amsterdam welfare system. The tool sought to incorporate good responsible AI practices, including incorporating community voices and technical fairness approaches. However, the developers could not debias the tool and each fairness approach revealed its normative consequences through having discriminatory outcomes for some protected category, with further challenges as intersectionality was considered (Guo et al., 2025). Coming to terms with this reality, the city abandoned the tool. Critics noted that the issue was pursuing AI for the purpose of welfare allocation and asked: Why not create a tool to identify people who could benefit from welfare but are outside the system? Or even, why not use the funds to tackle issues that keep people in a constant welfare state?

8.2.7. A call for new forms of creative, process-oriented innovation

Various questions arose at different points in this thesis: Should women get more loans (e.g., parity in true positive results)? Should tools be gender-aware and/or gender-differentiated? These are normative and loaded questions. Instead of answering these questions from my particular positionality, my findings prompt new questions around how those questions should be answered and by whom. These questions are therefore about *processes,* as opposed to desired outputs or how the tools *should* behave. They include, for example: How might AI technologies be developed differently, including the voices of community members in informing and influencing normative decisions? How might community members have control and agency over algorithmic tools and outcomes?

Instead of focusing on technical innovations, my findings call for innovations around *processes* which allow for more voices and agency to be included in normative decision-making and value-laden judgments in technology design and management. I further explore this in the future research agenda section. While not easy nor straightforward, as illustrated in other examples of community-informed governance approaches (D. Robinson, 2022), situated priorities of target users, including vulnerable communities, must be included. There are inherent tensions when private outside actors are the owners managing and delivering tools, particularly in the absence of regulation. Still, there can be solutions in private sector applications, though it requires reflexivity, grappling with tensions posed by existing organizational priorities, and funders supporting the pursuit of justice and equity. Unfortunately, market and profit priorities don't necessarily reward deeply engaging with and centering the voices and agency of community members. The sector needs new incentives – including regulation and funding for processes (as opposed to simply outputs or outcomes). It also needs new players that have different priorities, such as non-profits and civil society actors. Relatedly, solutions and approaches must reflect on what problems are being solved for and if they are avoiding or merely acting as band-aids to other core problems reinforcing a “politics of avoidance” (Shipton & Vitale, 2024).

8.3. Thesis contributions

My thesis provides valuable contributions to empirical literature, theory, methodologies for AI assessments, practice, and policy.

AI systems are increasingly governing different aspects of our lives. As the use of AI continues to grow in alternative lending, understanding the impacts of these systems is paramount. This includes understanding how they are impacted by and impact gender norms and inequalities. Overall, this thesis provides an important contribution in understanding the underlying logics, design, and management of ML-based credit assessment tools, as well as the gender implications. Using rare empirical interview data and survey data, I answer my overarching research question. I argued that underlying logics related to American banking and Silicon Valley result in gender “neutral” technologies that, in reality, learn from and embed existing gender economic inequalities. Institutional, economic, and development priorities skew algorithmic design and management decisions towards profit and scale, which result in particular design and management choices that have gender implications, expose potential discriminatory effects, and legitimize gender differences as rational. Gender differences are due to supply-side and user-side reasons, including women more often needing help and having challenges downloading and using the apps.

This is the first in-depth, critical study examining the underlying logics of ML-based credit assessment tools in LMICs, as well as one of the first on the gender impacts of an “AI for Good” application. The thesis therefore makes a direct contribution to both the literature on the use of AI in credit assessment, as well as the use of AI to support the SDGs. It provides important insights for various stakeholders on the implications of gender “neutral” or “blind” applications of ML (within and beyond ML-based alternative lending), and provides insights towards new pathways for inclusive and equitable innovation. It applies novel data and provides critical insight in the growing area of empirical AI ethics.

My results inform and support theory refinement. In particular, I expand existing theories of algorithmic bias and gender equity to offer encoded gender norms, which addresses how AI tools interact with gender and gender norms. While Ruha Benjamin provided default discrimination as a concept to particularly understand the impact of AI on race, with grounding in the West (Benjamin, 2019b), I am able to extend and explore the concept through both a gender lens and a focus on the majority world. This theoretical concept offers a way to understand how gender norms and inequalities are ingested and reproduced in algorithms informing decisions about humans that are developed under claims to be neutral. This concept connects to another concept I introduce, the objective algorithm paradox, whereby the belief in machine learning's objectivity results in being "blind" to certain demographics in ways that hide the gendered nature of algorithms and perpetuate discrimination. I also address a gap between SST (MacKenzie & Wajcman, 1999) and situated knowledges (Haraway, 1988) when assessing design and management of algorithmic technologies, to provide the concept of situated priorities, which posits that priorities are constructed based on specific contexts, pressures, and epistemological standpoints, leading to varied fairness practices with consequential algorithmic outcomes. The work thereby contributes to the feminist STS literature, as well as AI ethics literature broadly.

My thesis contributes to the methodological toolkit for academic AI evaluations and audits through presenting a sociotechnical, mixed-method, and grounded theory approach. This type of academic assessment is particularly suited for auditing algorithmic decision-making systems that are closed to public scrutiny (e.g., proprietary or opaque models), yet make decisions about people and have real implications for individuals and communities. By combining qualitative and quantitative techniques with a sociotechnical and feminist STS lens, this approach grappled with power dynamics and enabled deeper, contextualized understandings of AI systems and their implications, even in settings where access is limited. Collecting data for these types of studies is incredibly difficult, as studying developers and managers of AI systems is gate-kept and actors are often not incentivized to participate in academic studies, particularly given regulatory uncertainty. Admittedly, my positionality was important for me to access interviewees. Despite the challenges, this type of research is important and I hope my methods can offer learnings and inspiration to future researchers.

My findings have practical implications for fintechs, policymakers, and funders. I provide insight into the ways technologies can be more gender inclusive and equitable through concrete recommendations for fintechs in a later section of this conclusion. I present policy recommendations, highlighting that normative choices embedded in AI tools distributing resources require public governance and accountability. I also share recommendations for funders. Within this, I highlight the need for funding research exploring more equitable innovations that prioritize *processes* (not simply outputs or outcomes). I also call for research not just evaluating impacts among those who receive outputs from AI applications (e.g., loans via ML-based credit assessment apps), but assessing the models themselves. This comes at a particularly important time as research agendas in the international development space are predominantly focused on impact evaluations and not on assessing the models and how they are designed and managed, including underlying normative choices. While impact evaluations offer a helpful understanding of impacts among those who receive loans and are important, they are not sufficient as they do not capture who is deemed not creditworthy or "ground truths" in algorithmic systems. Holistic research agendas require both.

Together, this thesis has provided contributions both in the academic literature, theoretically, and practically. Section 8.5 of this conclusion details a future research agenda on gender and ML-based alternative lending with timely areas for research, including research I plan on undertaking through my own research agenda.

8.4. Limitations

There are various limitations of the research, which I outline here. I focus on high-level limitations and limitations present in my qualitative data. Additional, detailed limitations on the survey and quantitative analysis detailed in Chapter 7.

- **Gender binary:** While this research takes a gender-conscious approach and intentionally included nonbinary as an option in surveys and left interview questions about gender broad (e.g., not prompting "women" or "men"), all participants in the study identified within the gender binary or discussed gender in binary terms. As a result, the empirical data primarily reflects gendered experiences and impacts as they relate to dominant binary norms. This limits the ability of the study to speak directly to the experiences of nonbinary and gender-nonconforming individuals, who may face distinct challenges in algorithmic lending systems. However, the research does engage critically with how algorithmic tools reinforce or contest gender norms, which are themselves often constructed through binary logics. Future work would benefit from explicitly recruiting and centering the perspectives of nonbinary individuals, as well as expanding theoretical frameworks to better capture the multiplicity of gender identities and how they interact with automated decision-making systems.
- **Sample size of the survey**: The sample size for the survey data is on the smaller end. Though the survey data is valuable for identifying patterns, it should also be interpreted with caution given the relatively small sample size and because it is neither random nor representative of all app users. Results outlined in my Chapter 7 therefore may best be understood as exploratory, rather than representative. Future work could build on these findings with larger, more diverse samples to strengthen validity and have more robust statistical analysis.
- **Sample characteristics and self-selection in survey and interview data:** Both the interview and survey samples may be subject to self-selection bias. Participants opted into the study voluntarily, which may skew the sample towards individuals with particular experiences or interests. While the survey sample may not fully represent the broader population in Kenya who is using these apps, efforts were made to ensure diversity across key dimensions. Specifically, the sampling strategy aimed for gender balance (with a target of 50% women respondents), inclusion of first-time digital tool users, and representation of rural populations. Still, the survey remains

subject to self-selection bias and findings interpreted with caution when generalizing beyond the study sample.

- **Participant dynamics and honesty:** I rely on interviewees being honest, but they could have shared inaccuracies. In my first chapter, findings around gender differences in lending reported by fintechs rely on verbal estimates shared by interviewees. I did not receive any data on borrowers from fintechs themselves, despite my asking (and some agreeing to send data, but never following through). These estimates may reflect personal perceptions rather than verified records and could have been skewed – intentionally or not – to more positively reflect on the particular fintech or the industry broadly. That said, there was a consistency in the figures shared across interviewees from different fintechs within the same country contexts, which provides some triangulated confidence in the patterns reported. Nonetheless, the absence of direct, independently verified data remains a limitation. Relatedly, I also relied on survey participants' honesty and self-reported experiences. While the survey was anonymized to encourage candid responses, there's a possibility that some participants may have responded in ways they believed would be viewed favorably. This concern is particularly relevant given that the survey was distributed via a push notification from a fintech platform. Despite the clear consent language stating that the research was independent of the Lender, some respondents may have assumed a connection and offered more positive feedback in hopes of gaining favor or benefits. While this potential bias is difficult to know or quantify, it is important to acknowledge. Anonymity and independent ethical review mitigate this risk.
- **Lack of access to proprietary data and underlying models:** This research did not include access to actual borrower-level gender breakdowns or credit performance data from fintechs. Instead, I rely on interviewees' verbal reports of gender-related borrower trends and their interpretations of disparities in lending outcomes. While this approach allowed for valuable insider perspectives, it was shaped by data gatekeeping common in the fintech industry. This limits my ability to independently verify or analyze system-level outputs. As a result, gender

breakdowns reflect perceived patterns as opposed to empirical validation. While this limits my ability to independently verify claims, I do not see this as a methodological constraint. These accounts provide insight into how fintech actors understand (or not) gender dynamics in their systems.

- **Geographic breadth and strategic scope:** The interview data in this study includes participants across multiple country contexts, providing a broad, comparative view of ML-based credit assessment tools. However, this comes at the expense of depth in any one country-level or local setting. Therefore, the research does not offer deeply contextualized analyses of how country-specific legal, cultural, or economic conditions shape these algorithmic decision-making tools. This breadth was a deliberate choice. Given the overall scarcity of academic studies examining ML-based credit assessment tools themselves (including potential gender biases therein) and the well-known, significant access challenges involved in securing interviews with owners of proprietary AI tools, a wide sample inclusion strategy enhanced my ability to gather rare, firsthand data and insights. This approach supports my intention of this research as a foundational study. It also helps to map the industry and surface areas for more context-specific future research.

8.5. Agenda for Future Research

I present an agenda for future research with several pillars: (1) equitable technology innovation; (2) expansion of existing impact evaluations of ML-based credit assessment; and (3) other social impact and ethics issues presented by ML-based credit assessment. I then discuss a growing area of research and investment – gender-differentiated credit scoring – and outline research needed on this technology.

8.5.1. Research on equitable technology development

Fortunately, because algorithms are tools and instruments of values, they can be developed and employed in different ways in lending – including towards more equitable paradigms. This research stream recognizes that there are opportunities for adapting algorithmic lending tools to be more equitable, while also exploring how the tools can be leveraged to reflect on and challenge power hierarchies. This research stream delves into the implications of *who* develops technologies, why that matters, and how tradeoffs are navigated in value-laden algorithmic spaces. This research includes:

- **Equitable processes and co-creation**: Conduct participatory design and co-creation research for ML technologies with, for, and by vulnerable communities (e.g., women's savings and credit cooperatives). This research can examine methodologies and implications for integrating voices and agency of communities in normative decision-making, algorithmic design and management, while surfacing and exploring how tradeoffs are navigated. This research can include how different people perceive fairness and identify different ways that algorithms may behave and be evaluated to reflect different notions of fairness. Importantly, participatory design and co-creation research can be co-opted to serve the interests of people in power, and therefore must be done carefully and in ways of not imposing one's own agenda (Parpart et al., 2000). This research topic is on my personal research agenda.
- **Alternative models of AI development and ownership:** Investigate how the priorities and outcomes of algorithmic lending tools vary depending on who builds them (e.g., fintech startups, public institutions, women's savings and credit cooperatives, MFIs). This includes examining the influence of institutional values on tradeoff decisions and system transparency.
- **AI for financial accountability and systemic bias detection:** Research could support the development of AI tools designed specifically to audit or flag bias within existing financial systems, shifting AI's role from a mechanism of enforcement to one of oversight and reform.

8.5.2. Expanding existing impact evaluations of ML-based credit assessment tools

Existing impact evaluations of algorithmic decision-making systems often treat gender as a monolith and rely heavily on RCTs as the gold standard for evidence. While RCTs can offer valuable insights –

particularly on measurable outcomes – they are limited in what they reveal about the structural and sociotechnical dimensions of algorithmic bias. There are several research recommendations for studies in this space:

- **Incorporate intersectionality** to understand how algorithmic systems differently affect people across axes of race, class, gender identity, ability, and other dimensions of marginalization.
- **Recognize the epistemic limits of RCTs**, particularly their inability to surface issues embedded in model design and management – such as who is deemed creditworthy (or not), how data privacy is handled, surveillance implications, and broader accountability.
- **Balance methodological rigor with critical inquiry**, acknowledging the value of quantitative approaches while also integrating qualitative, participatory methods that can uncover harms not visible through metrics alone.

While impact evaluations and RCTs remain useful tools, a more expansive, reflexive approach is necessary to understand and challenge the power dynamics and structural biases in algorithmic systems, especially those shaping life chances under the guise of neutrality.

8.5.3. Expand the scope of research on ML-based credit assessment tools

Another key area of this research agenda is to broaden and deepen empirical and critical inquiry into ML-based credit assessment systems. This research focused on the industry broadly with interviewees spanning different global locations. Future research could focus more narrowly. Proposed directions for future research include:

- **Context-sensitive evaluations of specific tools:** Move beyond abstract assessments to conduct deeper, tool-specific audits. This includes examining how different models function in practice and how outcomes vary across demographic groups and geographies. This depends on partnership with a fintech to gain access to underlying models and data.

- **Cross-national comparisons:** Quantitatively investigate how gender disparities in algorithmic credit scoring differ across countries, and explore how these disparities correlate with broader structural factors, such as national indices of gender inequality.
- **Responsible AI dimensions:** This research focused on questions of fairness and bias in particular. However, there are other key areas that require further research. In particular, this includes building out research on the impacts of data privacy and increasing surveillance in credit assessment tools. As the tools continue to use as much data as possible and collect as much data as possible on people, they are practicing surveillance. More research is needed on the extent and implications of this surveillance, and how data is being used, repurposed, resold, and managed. Furthermore, with ongoing and rapid AI advancement, ML-based credit assessment tools are incorporating other advanced technologies such as generative AI. Research should assess the implications of these tech stacks and the risks they pose to vulnerable groups, using a gender lens and incorporating intersectionality.

8.5.4. Research on gender-differentiated algorithms

There are academics developing gender-differentiated algorithms, which mitigate some of the issues presented by gender "blind" tools. While making algorithms "gender aware" is appealing and seductive, it must be approached with extreme caution. These technologies can still learn from other (intersectional) disparities if not taken with caution. Furthermore, as gender-differentiated algorithm interventions are resulting in male-specific and female-specific credit scoring algorithms, they present two issues. First, they do not make space for non-binary individuals, potentially penalizing this already marginalized group. This echoes cautions from Sasha Costanza-Chock who highlights how transgender and non-binary people face unintended consequences by technologies that are built for the binary. Costanza-Chock reminds us that transgender and non-binary bodies can be flagged as "anomalous" by technologies with their gender non-conforming bodies regularly flagged as "risky" and/or invisibilized (Costanza-Chock, 2020). Secondly, these tools may inadvertently reward gendered behavior, thereby serving a policing function in how gender is practiced. The premise of this gender-differentiated algorithm intervention and associated research is normative, with various potential implications that must be unpacked as it continues to expand in LMICs and attract development funding, while advocating for changes in laws that allow for gender-differentiated lending algorithms. Research is needed not simply on the economic impacts of these technologies, including through RCTs, but on the models themselves, their socio-technical components and consequences, and their implications for how gender shapes and is shaped by the technology.

8.6. Recommendations

The following recommendations are directed at (1) fintechs, (2) investors and development funders, and (3) policymakers. Each stakeholder group plays a distinct but interconnected role to promote gender equitable access to finance and advance accountability in ML-based credit assessment in LMICs.

8.6.1. Recommendations for fintech companies

1. **Recognize the subjectivity of machine learning and embrace a mindset shift:** Fintech stakeholders should critically reflect on the values, assumptions, and priorities embedded in their models and decision-making processes. Rather than treating fairness as an objective technical outcome, they should acknowledge its normative dimensions and implications for gender equity. Then they should clearly articulate the rationale behind their chosen definitions, including by being transparent about tradeoffs and limitations, especially in relation to how different groups may be affected.
2. **Conduct internal and external audits:** Regular audits are important to evaluate model performance across different demographic groups and surface hidden disparities. These assessments should critically examine whether women (across intersecting identities) and other marginalized groups are being disproportionately denied loans or offered different terms.
3. **Partner with social scientists and critically engage with tradeoffs:** Collaborating with social scientists can help surface challenges and opportunities, while also informing limitations and

tradeoffs. Fintechs should not only identify tradeoffs among different priorities and considerations of fairness, but confront them head-on and transparently. At a higher level, this includes reflecting on how organizational values and incentive structures influence decisions.

4. **Practice transparency**: Even if developers cannot fully explain how an output was arrived at, there are ways to increase transparency and explainability of ML models. For example, developers can practice greater transparency and use research-backed resources to outline what data the ML models have been trained on (e.g., Datasheets for Datasets[27]) or to map out different aspects of ML models (e.g., Model Cards[28]). These forms of transparency are not currently standard practice in the fintech industry (or more broadly in the AI space) and without regulation or other incentives, is not necessarily in the best interest of the firms who are protecting their intellectual property and consider the models highly proprietary.[29] While different tools and resources exist in the ML space to increase transparency and explainability, they are not common practice amongst fintechs working in algorithmic-facilitated lending in LMICs (yet).
5. **Conduct gender-disaggregated UX design:** Gender-disaggregated UX design exercises can surface ways to make the app more approachable and usable for women, as well as for people broadly who are newer to digital financial apps and financial inclusion.
6. **Support financial literacy:** Provide simple, financial literacy educational tips on the app informed by a gender lens.

8.6.2. Recommendations for investors and development funders

These recommendations apply to all investors, including venture capital, private equity, and development funders. However, they may be particularly relevant for development funders, given their explicit social impact mandates.

[27] https://arxiv.org/abs/1803.09010
[28] https://dl.acm.org/doi/10.1145/3287560.3287596
[29] In certain countries and contexts this is changing. For example, the EU AI Act categorizes AI-based credit assessment tools in high risk categories which requires transparency and auditing.

1. **Reframe success – Shift from scale and efficiency to empowerment and equity:** A narrow focus on technology inclusion and scale can come at the cost of equity, empowerment, and accountability (O'Donnell & and Sweetman, 2018). Development funders in particular must be vigilant that their priorities and funding structures do not unintentionally incentivize fintechs to deprioritize or sideline gender and equity concerns. Instead of treating equity as an afterthought, funders should place it at the center, allocating resources to support equity-oriented design and implementation practices from the outset. This requires a mindset shift: from evaluating success primarily in terms of numbers reached and efficiency, to valuing impact on structurally excluded groups, even if those models are slower to – or never – scale.
2. **Support alternative organizations and models for more equitable technology futures:** Development funders and investors should expand support for organizations with alternative incentive structures (e.g., non-profits, microfinance organizations, cooperatives) that are better positioned to design AI technologies with, for, and by women and other marginalized groups. In the absence of strong regulation, investing in organizations that operate under different priorities and incentives offers a critical pathway to more equitable and accountable ML-based lending innovations. There is significant opportunity to foster research and innovation in this space, as participatory design and co-creation research is an area of growing interest in the AI ethics field. Therefore, supporting teams of practitioners and researchers is strongly advised.
3. **Fund processes, not simply outputs or outcomes:** A focus on measurable outputs or end results does not ensure that equity is meaningfully considered during development. Development funders should also invest in interventions focused specifically on processes and not require specific, up-front outputs or outcomes. It is within these processes that critical design decisions are made, shaping not only the technology itself but also whose needs it serves, whose perspectives are centered, and how power operates in its ongoing management. Prioritizing these processes is a critical aspect of data feminism (D'Ignazio & Klein, 2023).

4. **Insist on transparency:** Funders should require greater transparency from fintechs, including the use of tools like "model cards" and "datasheets for datasets" to document how models are built, what data is used, and how decisions are made. These tools help surface key assumptions, limitations, and potential biases. Requiring such tools can help normalize transparency practices across the sector without demanding firms disclose proprietary information.
5. **Provide gender technical advisory:** Alongside investment in the product and organization, funders should allocate resources to bring in gender experts and/or social scientists who can advise across the product lifecycle. This targeted support helps ensure that gender equity is not treated as an afterthought. This support must be accompanied by a clear commitment from the funder that gender considerations and equity is a priority.
6. **Fund research on inclusive governance of AI:** Support research that explores frameworks for oversight and regulation of ML-based alternative lending tools that center equity. This includes investigating models for public participation in algorithmic governance, mechanisms for transparency and accountability, and alternative approaches for community management and agency in ML-based decisions. As this is an emerging technology there is still much research needed, which is essential for building systems that are not only technically robust but also socially responsible.

8.6.3. Recommendations for policymakers

While the broader AI policy landscape remains fraught and in flux, there are actionable steps that can and must be taken now. I offer five policy recommendations. Among these, the most urgent is transparency, which is essential not only for regulatory oversight, but also for surfacing and addressing the normative assumptions and tradeoffs embedded in ML tools.

1. **Require transparency as a baseline policy measure:** This includes requiring regular reporting on model audits, disaggregated outcomes, and any disproportionate impacts on different

demographic groups. Within this, fintechs should report on fairness audits that go beyond statistical parity and incorporate intersectional analysis (e.g., gender and location or ethnicity). These audits should assess how different groups are affected by model decisions, not just overall error rates. Transparency requirements do not necessitate disclosing proprietary code or the models themselves which may be under IP, but should make visible the assumptions, tradeoffs, and impacts of ML systems on different populations. Transparency is a foundational step toward accountability and a necessary condition to identify unintended consequences and harms.

2. **Require justification and transparency of fairness approaches:** Rather than mandating a specific fairness metric, policymakers should require fintechs to clearly articulate how they define fairness in their models, which metrics or criteria they use, and why. This includes explaining tradeoffs, who is prioritized in the model's objectives, and what impacts are considered acceptable. By surfacing these normative choices, regulators can create space for public debate, scrutiny, and accountability, recognizing that fairness is not a fixed technical standard, but a contested social value that requires grappling by communities.
3. **Support independent monitoring and civil society engagement:** Fund and institutionalize independent bodies (e.g., academics, nonprofits, think tanks) that can assess and monitor algorithmic decision-making in credit assessment. This should include mechanisms for affected communities and civil society to participate in evaluating fairness claims and raising concerns.
4. **Build gender and AI expertise into regulatory agencies:** Regulators should be equipped with in-house staff or advisory capacity that brings both AI expertise and deep understanding of structural inequalities, particularly around gender, but also other aspects like ethnicity, ability, and caste (in the case of India). This can ensure that policy and oversight mechanisms reflect not only technical considerations, but social ones.
5. **Incentivize innovation in equitable ML practices:** This could include creating public funding mechanisms or regulatory sandboxes that reward or support AI innovations that have a focus on equity and participatory design.

8.7. Final reflections on my positionality

In the interviewees I saw a version of myself in the past: an eager development practitioner and entrepreneur frustrated by the status quo and driven to create a positive impact. I understand where they are coming from and I empathize with the choices and difficult tradeoffs they made and continue to be presented with. My positionality pushed me to a deeper reflection and prompted me to check any normative judgments that arose. I recognize how my positionality as a white, American-Canadian international development practitioner turned critical academic influenced my choices of theory and methodological approach. I deeply appreciate the probing questions from supervisors and critical readers that prompted me to ensure my analysis and generalizations remained true to my data. As I conclude this thesis, I recognize and hold space for the nuances that AI tools present. I push myself, academic colleagues, and international development peers to think critically and to grapple with the tradeoffs these technologies present in the face of increasing AI hype.

In the period that I'm writing this, my part-time position as the Gender & AI Fellow at USAID was eliminated as a result of the dismantling of USAID in the US, and I continue to separately facilitate a Gender and AI working group with several large international development donors. In the wake of USAID's elimination and as federal funding for international development has been near eliminated in the US, other national development funders continue facing budget cuts in an uncertain world. Meanwhile, US-based foundations are avoiding investments in gender-related projects for fear of being targeted by the federal administration and having their 501(c)3 tax status revoked. Furthermore, the Trump administration recently issued an executive order "preventing woke AI in the federal government", explicitly calling out diversity, equity, and inclusion as destructive ideologies (The White House, 2025). There is a clear retreat on topics of gender and race, while I see firsthand how the idea of AI to solve global problems more efficiently and productively amongst well-meaning practitioners is alluring and seductive. Yet, AI technologies must be approached with deep awareness and reflexivity. As in this thesis, I seek not to offer my own normative judgment on algorithmic tools with overly simple labels, but present the tradeoffs, illustrate tensions, and highlight promising paths forward that center opportunities to elevate people's voices and agency in algorithmic design and management.

8.8. Conclusion

As machine learning continues to proliferate globally, including in a variety of decision-making roles that affect people's access to resources and opportunities, understanding their gender and equity implications is critical. This thesis sought to unpack the "ground truth" in ML-based credit assessment models in LMICs to surface underlying mindsets, logics, and assumptions that inform how the technology is developed and managed. I surfaced how ML technologies embed, reflect, and advance deep beliefs and priorities, which have gender implications, including reinforcing gender inequities regarding access to finance and shaping gendered perceptions of creditworthiness. My findings emphasize that it matters who develops and designs ML technologies.

ML is a tool, neither good nor bad. Rather, one that is inherently sociotechnical: shaped by human values, institutional priorities, and the contexts in which it is developed and deployed. I leave this thesis optimistic, recognizing the myriad ways that technologies could be developed and managed in different ways, including in ways that center the agency of women and girls – as well as other marginalized communities. As the AI hype cycle continues, my hope is that researchers and developers do not get caught up in the seductive goals of speed and efficiency, but rather take a breath and reflect. Within this, we must ask whether we are being seduced into solving the wrong problems with this powerful technology, and instead explore how these tools might be repurposed to address different questions, serve different needs, and realize alternative, more equitable futures. This reframing enables expansive visioning and unlocks opportunities for new innovations, partnerships, and research questions.

As I bring this thesis to a close, I find myself drawn to this deeper invitation – not just to critique, but to imagine. To imagine what becomes possible when we step beyond unchecked beliefs and assumptions and toward new ways of asking and building – ways grounded in the perspectives and priorities of those historically excluded from technology design and production. I close my eyes and picture a door that this reframing opens. It has a heavy metal latch, which I pull up before walking under the curved wooden arch. I am greeted by jasmine bushes, their scent filling me with a familiar calm, and I see several dirt paths winding through birch trees. I choose one and begin to curve past the trees, finding myself amidst wildflowers – daisies, columbines, bluebells, and budding pink blooms unlike any I've seen before. I kneel beside one: a soft pink flower streaked with golden orange that seems to both capture and hold a whole universe in its layered petals. This imagined garden is not an escape, but a symbol of what might grow when we invite inclusive knowledge production and make space for care, plurality, and collective vision in the research and development of AI technologies.

REFERENCES

Adams, R. (2021). Can artificial intelligence be decolonized? *Interdisciplinary Science Reviews*, *46*(1–2), 176–197. https://doi.org/10.1080/03080188.2020.1840225

Adams, R., Weale, S., & Barr, C. (2020, August 13). A-level results: Almost 40% of teacher assessments in England downgraded. *The Guardian*. https://www.theguardian.com/education/2020/aug/13/almost-40-of-english-students-have-a-level-results-downgraded

Agarwal, B. (1995). *A Field of One's Own: Gender and Land Rights in South Asia*. Cambridge University Press. https://doi.org/10.1017/CBO9780511522000

AI for Good. (n.d.). AI for Good. Retrieved May 23, 2025, from https://aiforgood.itu.int/

AI for SDGs Observatory. (n.d.). Center for Long-Term Artificial Intelligence. Retrieved May 23, 2025, from https://ai-for-sdgs.academy/observatory

Akolgo, I. A. (2023). On the contradictions of Africa's fintech boom: Evidence from Ghana. *Review of International Political Economy*, *30*(5), 1639–1659. https://doi.org/10.1080/09692290.2023.2225142

Alcoff, L. (1988). Cultural Feminism versus Post-Structuralism: The Identity Crisis in Feminist Theory. *Signs: Journal of Women in Culture and Society*, *13*(3), 405–436. https://doi.org/10.1086/494426

Alkire, S. (2005). Why the Capability Approach? *Journal of Human Development*, *6*(1), 115–135. https://doi.org/10.1080/146498805200034275

Allen, F., Demirguc-Kunt, A., Klapper, L., & Martinez Peria, M. S. (2016). The foundations of financial inclusion: Understanding ownership and use of formal accounts. *Journal of Financial Intermediation*, *27*, 1–30. https://doi.org/10.1016/j.jfi.2015.12.003

Andreeva, G., & Matuszyk, A. (2019). The Law of Equal Opportunities or Unintended Consequences?: The Effect of Unisex Risk Assessment in Consumer Credit. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *182*(4), 1287–1311. https://doi.org/10.1111/rssa.12494

Angwin, J., & Larson, J. (2016, July 29). *ProPublica Responds to Company's Critique of Machine Bias Story*. ProPublica. https://www.propublica.org/article/propublica-responds-to-companys-critique-of-machine-bias-story

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). *Machine Bias*. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Armendariz de Aghion, B., & Morduch, J. (2010). *The Economics of Microfinance, Second Edition* [MIT Press Books]. The MIT Press. https://econpapers.repec.org/bookchap/mtptitles/0262513986.htm

Arraiz, I., Bruhn, M., & Stucchi, R. (2017). Psychometrics as a Tool to Improve Credit Information. *The World Bank Economic Review*, *30*(Supplement_1), S67–S76.

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The Miracle of Microfinance? Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics*, *7*(1), 22–53. https://doi.org/10.1257/app.20130533

Banerjee, A., Karlan, D., & Zinman, J. (2015). Six Randomized Evaluations of Microcredit: Introduction and Further Steps. *American Economic Journal: Applied Economics*, *7*(1), 1–21. https://doi.org/10.1257/app.20140287

Barabas, C., Doyle, C., Rubinovitz, J., & Dinakar, K. (2020). Studying up: Reorienting the study of algorithmic fairness around issues of power. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 167–176. https://doi.org/10.1145/3351095.3372859

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, *83*, 405–417. https://doi.org/10.1016/j.eswa.2017.04.006

Barik, K., & Misra, S. (2024). Analysis of customer reviews with an improved VADER lexicon classifier. *Journal of Big Data*, *11*(1), 10. https://doi.org/10.1186/s40537-023-00861-x

Barocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and Machine Learning: Limitations and Opportunities*. The MIT Press. https://mitpress.mit.edu/9780262048613/fairness-and-machine-learning/

Barocas, S., & Selbst, A. D. (2016). *Big Data's Disparate Impact* (SSRN Scholarly Paper 2477899). Social Science Research Network. https://doi.org/10.2139/ssrn.2477899

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2019). *Consumer-Lending Discrimination in the FinTech Era* (Working Paper 25943). National Bureau of Economic Research. https://doi.org/10.3386/w25943

Bauer, G. R. (2014). Incorporating intersectionality theory into population health research methodology: Challenges and the potential to advance health equity. *Social Science & Medicine*, *110*, 10–17. https://doi.org/10.1016/j.socscimed.2014.03.022

Beck, T., Demirguc-Kunt, A., & Honohan, P. (2009). Access to Financial Services: Measurement, Impact, and Policies. *The World Bank Research Observer*, *24*(1), 119–145.

Benjamin. (2019a). *Race After Technology: Abolitionist Tools for the New Jim Code | Wiley*. Polity. https://www.wiley.com/en-nl/Race+After+Technology%3A+Abolitionist+Tools+for+the+New+Jim+Code-p-9781509526437

Benjamin, R. (2019b). *Race After Technology: Abolitionist Tools for the New Jim Code | Wiley*. Wiley.Com. https://www.wiley.com/en-us/Race+After+Technology%3A+Abolitionist+Tools+for+the+New+Jim+Code-p-9781509526437

Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the Rise of FinTechs: Credit Scoring Using Digital Footprints. *The Review of Financial Studies*, *33*(7), 2845–2897.

Bevir, M., & Blakely, J. (2018). *Interpretive Social Science: An Anti-naturalist Approach*. Oxford University Press.

Bharadwaj, P., & Suri, T. (2020). Improving Financial Inclusion through Digital Savings and Credit. *AEA*

Papers and Proceedings, *110*, 584–588. https://doi.org/10.1257/pandp.20201084

Björkegren, D., Blumenstock, J., Folajimi-Senjobi, O., Mauro, J., & Nair, S. R. (2022). *Instant Loans Can Lift Subjective Well-Being: A Randomized Evaluation of Digital Credit in Nigeria* (arXiv:2202.13540). arXiv. https://doi.org/10.48550/arXiv.2202.13540

Björkegren, D., & Grissen, D. (2018). The Potential of Digital Credit to Bank the Poor. *AEA Papers and Proceedings*, *108*, 68–71. https://doi.org/10.1257/pandp.20181032

Björkegren, D., & Grissen, D. (2020). Behavior Revealed in Mobile Phone Usage Predicts Loan Repayment. *The World Bank Economic Review*, *34*(3), 618–634. https://doi.org/10.1093/wber/lhz006

Blanco-Oliver, A. J., Irimia-Diéguez, A. I., & Vázquez-Cueto, M. J. (2023). Is there an optimal microcredit size to maximize the social and financial efficiencies of microfinance institutions? *Research in International Business and Finance*, *65*, 101980. https://doi.org/10.1016/j.ribaf.2023.101980

Blanco-Oliver, A., Reguera-Alvarado, N., & Veronesi, G. (2021). Credit risk in the microfinance industry: The role of gender affinity. *Journal of Small Business Management*, *59*(2), 280–311.

Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation | The Journal of Machine Learning Research. *The Journal of Machine Learning Research*, *3*, 993–1022.

Bono, T., Croxson, K., & Giles, A. (2021). Algorithmic fairness in credit scoring. *Oxford Review of Economic Policy*, *37*(3), 585–617.

Brailovskaya, V., Dupas, P., & Robinson, J. (2021). *Is Digital Credit Filling a Hole or Digging a Hole? Evidence from Malawi* (SSRN Scholarly Paper 3991382). Social Science Research Network. https://papers.ssrn.com/abstract=3991382

Branch International | About. (n.d.). Branch. Retrieved May 19, 2025, from https://branch.co/about

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial

Gender Classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

Burnham, K., & Anderson, D. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic*. Springer. https://link.springer.com/book/10.1007/b97636

Burrell, J. (2024). Automated decision-making as domination. *First Monday*. https://doi.org/10.5210/fm.v29i4.13630

Butler, J. (2011). *Bodies That Matter: On the Discursive Limits of Sex*. Routledge. https://doi.org/10.4324/9780203828274

Buvinic, M., & Gokhroo, S. (2023). A NARRATIVE REVIEW IN LOW- AND MIDDLE-INCOME COUNTRIES. *Center for Global Development, Working Papers*.

Buvinic, M., & Ruf, R. (2022, June 14). *Towards Women's Financial Inclusion: Gender Data Diagnostics of Six Countries - Data2X*. Data2X. https://data2x.org/resource-center/wfid-synthesis-brief/, https://data2x.org/resource-center/wfid-synthesis-brief/

Cao, L., Yang, Q., & Yu, P. S. (2021). Data science and AI in FinTech: An overview. *International Journal of Data Science and Analytics*, *12*(2), 81–99. https://doi.org/10.1007/s41060-021-00278-w

Caton, S., & Haas, C. (2024). Fairness in Machine Learning: A Survey. *ACM Comput. Surv.*, *56*(7), 166:1-166:38. https://doi.org/10.1145/3616865

CEGA. (2024). *Mobile Instant Credit: Impacts, Challenges, and Lessons for Consumer Protection*. Center for Effective Global Action. https://cega.berkeley.edu/document/mobile-instant-credit-impacts-challenges-and-lessons-for-consumer-protection/

Central Bank of Kenya. (2024). *2024 FinAcess Household Survey Report*. Central Bank of Kenya. https://www.knbs.or.ke/reports/2024-finacess-household-survey-report/

CGAP. (n.d.). *Digital Credit Survey*. CGAP. Retrieved June 7, 2025, from https://www.cgap.org/topics/collections/digital-credit

Chakrabarti, P., & Frye, M. (2017). A mixed-methods framework for analyzing text data: Integrating computational techniques with qualitative methods in demography. *Demographic Research*, *37*, 1351–1382. https://doi.org/10.4054/DemRes.2017.37.42

Chalwe-Mulenga, M., & Duflos, E. (2021, October 19). *The Evolving Nature and Scale of Consumer Risks in Digital Finance*. CGAP. https://www.cgap.org/blog/evolving-nature-and-scale-of-consumer-risks-in-digital-finance, https://www.cgap.org/blog/evolving-nature-and-scale-of-consumer-risks-in-digital-finance

Charmaz, K. (2024). *Constructing Grounded Theory*. SAGE Publications. https://collegepublishing.sagepub.com/products/constructing-grounded-theory-3-255601

Chen, A. Y., Even-Tov, O., Kang, J. K., & Wittenberg-Moerman, R. (2025). Digital Lending and Financial Well-Being: Through the Lens of Mobile Phone Data. *The Accounting Review*, 1–25. https://doi.org/10.2308/TAR-2024-0046

Chen, S., Doerr, S., Frost, J., Gambacorta, L., & Shin, H. S. (2023). The fintech gender gap. *Journal of Financial Intermediation*, *54*, 101026. https://doi.org/10.1016/j.jfi.2023.101026

Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications*, *10*(1), 567. https://doi.org/10.1057/s41599-023-02079-x

Cheong, M., Abedin, E., Ferreira, M., Reimann, R., Chalson, S., Robinson, P., Byrne, J., Ruppanner, L., Alfano, M., & Klein, C. (2024). Investigating Gender and Racial Biases in DALL-E Mini Images. *ACM J. Responsib. Comput.*, *1*(2), 13:1-13:20. https://doi.org/10.1145/3649883

Chioda, L., Gertler, P., Higgins, S., & Medina, P. C. (2024). *FinTech Lending to Borrowers with No Credit History* (Working Paper 33208). National Bureau of Economic Research. https://doi.org/10.3386/w33208

Chlouverakis, K. (2024). *How artificial intelligence is reshaping the financial services industry*. EY. https://www.ey.com/en_gr/insights/financial-services/how-artificial-intelligence-is-reshaping-the-financial-services-industry

Christin, A. (2020). The ethnographer and the algorithm: Beyond the black box. *Theory and Society*, *49*(5), 897–918. https://doi.org/10.1007/s11186-020-09411-3

Chu, Y., Sun, C., Zhang, B., & Zhao, D. (2023). *Fintech and Gender Discrimination* (SSRN Scholarly Paper 4322257). Social Science Research Network. https://doi.org/10.2139/ssrn.4322257

Citron, D., & Pasquale, F. (2014). The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, *89*(1), 1.

Clouser, K. D., & Gert, B. (1990). A critique of principlism. *The Journal of Medicine and Philosophy*, *15*(2), 219–236. https://doi.org/10.1093/jmp/15.2.219

Collett, C. (2024). The Hustle: How Struggling to Access Elites for Qualitative Interviews Alters Research and the Researcher. *Qualitative Inquiry*, *30*(7), 555–567. https://doi.org/10.1177/10778004231188054

Collins, P. H. (2009). *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge & CRC Press. https://www.routledge.com/Black-Feminist-Thought-Knowledge-Consciousness-and-the-Politics-of-Empowerment/Hill-Collins/p/book/9780415964722

Cooper, D., Haley, B., Govender, D., Mokdad, A., & Mkhize, M. (2025, April 2). Kenya's AI Strategy 2025–2030: Signals for Global Companies Operating in Africa. *Global Policy Watch*. https://www.globalpolicywatch.com/2025/04/kenyas-ai-strategy-2025-2030-signals-for-global-companies-operating-in-africa/

Corbett-Davies, S., Gaebler, J. D., Nilforoshan, H., Shroff, R., & Goel, S. (2023). *The Measure and Mismeasure of Fairness* (arXiv:1808.00023). arXiv. https://doi.org/10.48550/arXiv.1808.00023

Corrales-Barquero, R., Marín-Raventós, G., & Barrantes, E. G. (2021). A Review of Gender Bias Mitigation in Credit Scoring Models. *2021 Ethics and Explainability for Responsible Data Science (EE-RDS)*, 1–10. https://doi.org/10.1109/EE-RDS53766.2021.9708589

Costanza-Chock, S. (2020). *Design Justice*. The MIT Press. https://mitpress.mit.edu/9780262043458/design-justice/

Cowgill, B., Dell'Acqua, F., Deng, S., Hsu, D., Verma, N., & Chaintreau, A. (2020). *Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics* (arXiv:2012.02394). arXiv. https://doi.org/10.48550/arXiv.2012.02394

Credit experts since 1899. (n.d.). Equifax. Retrieved May 16, 2025, from https://www.equifax.co.uk/resources/what-we-do/credit-experts-since-1899.html

Crenshaw, K., Gotanda, N., Peller, G., & Thomas, K. (1996). *Critical Race Theory*. The New Press. https://thenewpress.com/books/critical-race-theory

Criado, N., Ferrer, X., & Such, J. M. (2021). Attesting Digital Discrimination Using Norms. *International Journal of Interactive Multimedia and Artificial Intelligence*, *6*(Special Issue on Artificial Intelligence, Paving the Way to the Future), 16–23.

Criado-Perez, C. (2021). *Invisible Women: Data Bias in a World Designed for Men|Paperback*. Abrams Press. https://www.barnesandnoble.com/w/invisible-women-caroline-criado-perez/1129490658

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, *47*(2), 448–474. https://doi.org/10.1257/jel.47.2.448

Dahal, M., & Fiala, N. (2020). What do we know about the impact of microfinance? The problems of statistical power and precision. *World Development*, *128*, 104773. https://doi.org/10.1016/j.worlddev.2019.104773

Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, *91*, 106263. https://doi.org/10.1016/j.asoc.2020.106263

Dastin, J. (2018, October 11). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/

Daston, L., & Galison, P. (2007). *Objectivity*. https://press.princeton.edu/books/paperback/9781890951795/objectivity de Andrés, P., Gimeno, R., & Mateos de Cabo, R. (2021). The gender gap in bank credit access. *Journal of Corporate Finance*, *71*, 101782. https://doi.org/10.1016/j.jcorpfin.2020.101782

De, S., Jangra, S., Agarwal, V., Johnson, J., & Sastry, N. (2023). Biases and Ethical Considerations for Machine Learning Pipelines in the Computational Social Sciences. In A. Mukherjee, J. Kulshrestha, A. Chakraborty, & S. Kumar (Eds.), *Ethics in Artificial Intelligence: Bias, Fairness and Beyond* (pp. 99–113). Springer Nature. https://doi.org/10.1007/978-981-99-7184-8_6

Demirguc-Kunt, A., Klapper, L., & Singer, D. (2013). *Financial Inclusion and Legal Discrimination against Women: Evidence from Developing Countries* (Policy Research Working Papers). World Bank. https://www.researchgate.net/publication/381700880_Financial_Inclusion_and_Legal_Discrimination_against_Women_Evidence_from_Developing_Countries

Demirgüç-Kunt, A., & Singer, D. (2017). *Financial Inclusion and Inclusive Growth: A Review of Recent Empirical Evidence* (SSRN Scholarly Paper 2958542). Social Science Research Network. https://papers.ssrn.com/abstract=2958542

Denzin, N., & Lincoln, Y. (2017). *The SAGE Handbook of Qualitative Research*. SAGE Publications. https://us.sagepub.com/en-us/nam/the-sage-handbook-of-qualitative-research/book242504

Deshpande, R., & Koning, A. (2023). *The Impact of Financial Inclusion on Young Women's Well-being: A Survey of Evidence and Recommendations for Practitioners*. CGAP. https://www.cgap.org/research/publication/impact-of-financial-inclusion-on-young-womens-well-being-survey-of-evidence, https://www.cgap.org/research/publication/impact-of-financial-inclusion-on-young-womens-well-being-survey-of-evidence

D'Espallier, B., Guérin, I., & Mersland, R. (2011). Women and Repayment in Microfinance: A Global Analysis. *World Development*, *39*(5), 758–772. https://doi.org/10.1016/j.worlddev.2010.10.008

Di Maggio, M., & Ratnadiwakara, D. (2022). *Invisible Primes: Fintech Lending with Alternative Data* (SSRN Scholarly Paper 3937438). https://doi.org/10.2139/ssrn.3937438

D'Ignazio, C., & Klein, L. (2023). *Data Feminism*. The MIT Press. https://mitpress.mit.edu/9780262547185/data-feminism/

Doyal, L. (2000). Gender equity in health: Debates and dilemmas. *Social Science & Medicine*, *51*(6), 931–939. https://doi.org/10.1016/S0277-9536(00)00072-1

Edmonds, W. A., & Kennedy, T. D. (2017). Convergent-Parallel Approach. In *An Applied Guide to Research Designs:An Applied Guide to Research Designs: Quantitative, Qualitative, and Mixed Methods*. SAGE Publications. https://methods.sagepub.com/book/mono/an-applied-guide-to-research-designs-2e/chpt/convergentparallel-approach

Elson, D. (1995). Gender Awareness in Modeling Structural Adjustment. *World Development*, *23*(11), 1851–1868. https://doi.org/10.1016/0305-750X(95)00087-S

Elson, D. (1999). Labor Markets as Gendered Institutions: Equality, Efficiency and Empowerment Issues. *World Development*, *27*(3), 611–627. https://doi.org/10.1016/S0305-750X(98)00147-8

Empson, L. (2018). Elite interviewing in professional organizations. *Journal of Professions and Organization*, *5*(1), 58–69. https://doi.org/10.1093/jpo/jox010

Eubanks, V. (2018a). *Automating Inequality*. St. Martin's Press. https://us.macmillan.com/books/9781250074317/automatinginequality/

Eubanks, V. (2018b). *Automating Inequality*. Macmillan Publishers. https://us.macmillan.com/books/9781250074317/automatinginequality

Fan, Y., Yang, Y., & Qin, Y. (2013). Credit scoring model based on PCA and improved tree augmented Bayesian Classification. *IET International Conference on Information and Communications Technologies (IETICT 2013)*, 169–175. https://doi.org/10.1049/cp.2013.0051

Faulkner, W. (2001). The technology question in feminism: A view from feminist technology studies. *Women's Studies International Forum*, *24*(1), 79–95. https://doi.org/10.1016/S0277-5395(00)00166-7

Federal Reserve Board. (2007). *Report to the Congress on Credit Scoring and Its Effects on the*

Availability and Affordability of Credit. Federal Reserve. https://www.federalreserve.gov/boarddocs/rptcongress/creditscore/general.htm

Ferrer, X., Nuenen, T. van, Such, J. M., Coté, M., & Criado, N. (2021). Bias and Discrimination in AI: A Cross-Disciplinary Perspective. *IEEE Technology and Society Magazine*, *40*(2), 72–80. https://doi.org/10.1109/MTS.2021.3056293

FICO. (2025, February 18). *TransUnion and FICO Partner to Introduce Groundbreaking Risk Solutions to Kenya to Expand Credit Access*. FICO. https://www.fico.com/en/newsroom/transunion-and-fico-partner-introduce-groundbreaking-risk-solutions-kenya-expand-credit

Findexable. (2021). *Fintech Diversity Radar: Diversity for Growth*. https://findexable.com/fintech-diversity-radar-fdr/

Fleisig, E., Smith, G., Bossi, M., Rustagi, I., Yin, X., & Klein, D. (2024). Linguistic Bias in ChatGPT: Language Models Reinforce Dialect Discrimination. In Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (Eds.), *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 13541–13564). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.750

Foucault, M. (2006). *Madness and Civilization*. Routledge. https://www.routledge.com/Madness-and-Civilization/Foucault/p/book/9780415253857

Foucault, M. (2020). *Power*. Penguin. https://www.penguin.co.uk/books/23075/power-by-foucault-michel/9780241435083

Fountain, J. (2001). *Building the Virtual State*. Brookings Institution Press. https://www.brookings.edu/books/building-the-virtual-state/

Fourcade, M., & Healy, K. (2024). *The Ordinal Society*. Harvard University Press. https://www.hup.harvard.edu/books/9780674971141

Freelon, D. (2014). On the Interpretation of Digital Trace Data in Communication and Social Computing Research. *Journal of Broadcasting & Electronic Media*, *58*(1), 59–75.

https://doi.org/10.1080/08838151.2013.875018

Friederici, N., Ojanperä, S., & Graham, M. (2017). The Impact of Connectivity in Africa: Grand Visions and the Mirage of Inclusive Digital Development. *THE ELECTRONIC JOURNAL OF INFORMATION SYSTEMS IN DEVELOPING COUNTRIES*, *79*(1), 1–20. https://doi.org/10.1002/j.1681-4835.2017.tb00578.x

Friedler, S. A., Scheidegger, C., Venkatasubramanian, S., Choudhary, S., Hamilton, E. P., & Roth, D. (2018). *A comparative study of fairness-enhancing interventions in machine learning* (arXiv:1802.04422). arXiv. https://doi.org/10.48550/arXiv.1802.04422

Friedman, B., & Nissenbaum, H. (2017). Bias in Computer Systems. In *Computer Ethics* (1st ed., pp. 215–232). Routledge. https://www.researchgate.net/publication/329747893_Bias_in_Computer_Systems

FT Partners. (2022). *Women in Fintech*. https://www.ftpartners.com/fintech-research/women-in-fintech2022

Fuster, A., Goldsmith-Pinkham, P. S., Ramadorai, T., & Walther, A. (2021). *Predictably Unequal? The Effects of Machine Learning on Credit Markets* (SSRN Scholarly Paper 3072038). Social Science Research Network. https://doi.org/10.2139/ssrn.3072038

Gabor, D., & Brooks, S. (2017). The digital revolution in financial inclusion: International development in the fintech era. *New Political Economy*, *22*(4), 423–436.

Gándara, D., Anahideh, H., Ison, M. P., & Picchiarini, L. (2024). *Inside the Black Box: Detecting and Mitigating Algorithmic Bias across Racialized Groups in College Student-Success Prediction* (arXiv:2301.03784). arXiv. https://doi.org/10.48550/arXiv.2301.03784

Gandhi, L. (2019). *Postcolonial Theory: A Critical Introduction: Second Edition, second edition* (p. 296 Pages). Columbia University Press.

Garcia, A. C. B., Garcia, M. G. P., & Rigobon, R. (2024). Algorithmic discrimination in the credit domain: What do we know about it? *AI & SOCIETY*, *39*(4), 2059–2098. https://doi.org/10.1007/s00146-023-01676-3

Gender discrimination in credit access prohibited. (2023). Our World in Data. https://ourworldindata.org/grapher/discrimination-access-to-credit-gender

Gerson, K. (2020). Analyzing Interviews: Making Sense of Complex Material. In K. Gerson & S. Damaske (Eds.), *The Science and Art of Interviewing* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780199324286.003.0006

Girrbach, L., Alaniz, S., Smith, G., & Akata, Z. (2025). *A Large Scale Analysis of Gender Biases in Text-to-Image Generative Models* (arXiv:2503.23398). arXiv. https://doi.org/10.48550/arXiv.2503.23398

Golden, A. (2024, September 30). *Are FICO Scores the same around the world?* FICO. https://www.fico.com/blogs/are-fico-scores-same-around-world

Golder, S., & Macy, M. (2014). Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology*, *40*, 129–152.

Grand View Research. (2023). *Global Alternative Financing Market Size & Share Report, 2023-2030*. Grand View Research. https://www.grandviewresearch.com/industry-analysis/alternative-financing-market-report

Gryz, J., & Rojszczak, M. (2021). Black box algorithms and the rights of individuals: No easy solution to the "explainability" problem. *Internet Policy Review*, *10*(2). https://policyreview.info/articles/analysis/black-box-algorithms-and-rights-individuals-no-easy-solution-explainability

GSMA. (2024). *The Mobile Gender Gap Report*. GSMA. https://www.gsma.com/r/gender-gap/

GSMA. (2025). *The Mobile Gender Gap Report*. GSMA. https://www.gsma.com/r/gender-gap/

Guo, E., Geiger, G., & Braun, J.-C. (2025). *Inside Amsterdam's high-stakes experiment to create fair welfare AI*. MIT Technology Review. https://www.technologyreview.com/2025/06/11/1118233/amsterdam-fair-welfare-ai-discriminatory-algorithms-failure/

Gutierrez, M. (2021). New Feminist Studies in Audiovisual Industries| Algorithmic Gender Bias and

Audiovisual Data: A Research Agenda. *International Journal of Communication*, *15*(0), Article 0.

Ha, D., Le, P., & Nguyen, D. K. (2025). Financial inclusion and fintech: A state-of-the-art systematic literature review. *Financial Innovation*, *11*(1), 69. https://doi.org/10.1186/s40854-024-00741-0

Haas, C. (2019). The Price of Fairness—A Framework to Explore Trade-Offs in Algorithmic Fairness. *ICIS 2019 Proceedings*. https://aisel.aisnet.org/icis2019/data_science/data_science/19

Haenssgen, M. J., & and Ariana, P. (2018). The place of technology in the Capability Approach. *Oxford Development Studies*, *46*(1), 98–112. https://doi.org/10.1080/13600818.2017.1325456

Hall, P., & Ellis, D. (2023). A systematic review of socio-technical gender bias in AI algorithms. *Online Information Review*, *47*(7), 1264–1279. https://doi.org/10.1108/OIR-08-2021-0452

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, *14*(3), 575–599. https://doi.org/10.2307/3178066

Harding, S. (1995). "Strong objectivity": A response to the new objectivity question. *Synthese*, *104*(3), 331–349. https://doi.org/10.1007/BF01064504

Hardt, M., Price, E., & Srebro, N. (2016). *Equality of Opportunity in Supervised Learning* (arXiv:1610.02413). arXiv. https://doi.org/10.48550/arXiv.1610.02413

Hasler, A., & Lusardi, A. (2017). *The Gender Gap in Financial Literacy: A Global Perspective*. George Washington University. https://gflec.org/research/

Hassan, Y. (2023). Governing algorithms from the South: A case study of AI development in Africa. *AI & SOCIETY*, *38*(4), 1429–1442. https://doi.org/10.1007/s00146-022-01527-7

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (pp. xvii, 507). The Guilford Press.

Herrerias, R., & M. O. Alvarez, C. (2023). *Financial Behavior and Degrees of Financial Inclusion* (SSRN Scholarly Paper 3717100). Social Science Research Network. https://doi.org/10.2139/ssrn.3717100

Hesse-Biber, S., & Griffin, A. J. (2015). Feminist Approaches to Multimethod and Mixed Methods Research: Theory and Praxis. In S. N. Hesse-Biber & R. B. Johnson (Eds.), *The Oxford*

Handbook of Multimethod and Mixed Methods Research Inquiry (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199933624.013.6

Hesse-Biber, S., LinaLeavy, P., & Leckenby, D. (2007). Feminist Research Practice—Feminist Approaches to Mixed-Methods Research. In *Sage Research Methods*. Sage. https://methods.sagepub.com/book/mono/feminist-research-practice/chpt/feminist-approaches-mixedmethods-research

Hohnen, P., Ulfstjerne, M. A., & Krabbe, M. S. (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), Article 1. https://doi.org/10.5617/jea.8315

Holdsworth, J., & Scapicchio, M. (2024, June 17). *What Is Deep Learning?* https://www.ibm.com/think/topics/deep-learning

Holstein, K., Wortman Vaughan, J., Daumé, H., Dudik, M., & Wallach, H. (2019). Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3290605.3300830

hooks, B. (1984). *Feminist Theory: From Margin to Center*. https://bellhooksbooks.com/product/feminist-theory/

Horton, L. (2017). *Women and Microfinance in the Global South: Empowerment and Disempowerment Outcomes*. Cambridge University Press. https://digitalcommons.chapman.edu/sociology_books/17

Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied Logistic Regression, 3rd Edition*. Wiley. https://www.wiley.com/en-us/Applied+Logistic+Regression%2C+3rd+Edition-p-9780470582473

Hsu, J. (2018). *Strava Data Heat Maps Expose Military Base Locations Around the World | WIRED*. Wired. https://www.wired.com/story/strava-heat-map-military-bases-fitness-trackers-privacy/

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), Article 1. https://doi.org/10.1609/icwsm.v8i1.14550

ILO. (2024). *The Impact of Care Responsibilities on Women's Labour Force Participation*. International Labour Organization (ILO). https://www.ilo.org/resource/news/unpaid-care-work-prevents-708-million-women-participating-labour-market

Inclusive Fintech 50. (n.d.). Center for Financial Inclusion. Retrieved December 7, 2024, from https://inclusivefintech50.com/research/fintech/inclusive-fintech-50/

InsightAce Analytic. (2024). *AI In The Credit-Scoring Market Latest Trends Analysis Report in 2024*. InsightAce Analytic. https://www.insightaceanalytic.com/report/ai-in-the-credit-scoring-market/2578?utm_source=whatech&utm_medium=refferal&utm_campaign=shorturl&utm_content=whatech-com-902002

International Finance Corporation. (2024). *Her Fintech Edge: Market Insights for Inclusive Growth* [Text/HTML]. IFC. https://www.ifc.org/en/insights-reports/2024/her-fintech-edge-market-insights-for-inclusive-growth

Irani, L., Vertesi, J., Dourish, P., Philip, K., & Grinter, R. E. (2010). Postcolonial computing: A lens on design and development. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1311–1320. https://doi.org/10.1145/1753326.1753522

Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Financial Management*, *48*(4), 1009–1029. https://doi.org/10.1111/fima.12295

Jain, S., & Gabor, D. (2020). The Rise of Digital Financialisation: The Case of India. *New Political Economy*, *25*(5), 813–828. https://doi.org/10.1080/13563467.2019.1708879

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications*, *78*(11), 15169–15211. https://doi.org/10.1007/s11042-018-6894-4

Johnen, C., & Mußhoff, O. (2023). Digital credit and the gender gap in financial inclusion: Empirical evidence from Kenya. *Journal of International Development*, *35*(2), 272–295. https://doi.org/10.1002/jid.3687

Johnson, K., Pasquale, F., & Chapman, J. (2019). Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation. *Fordham Law Review*, *88*(2), 499.

J-PAL. (n.d.). *Examples of survey questions related to women's empowerment*. Retrieved June 7, 2025, from https://www.povertyactionlab.org/page/practical-guide-measuring-womens-and-girls-empowerment-impact-evaluations

J-PAL. (2023, May 1). *Microcredit: Impacts and promising innovations*. The Abdul Latif Jameel Poverty Action Lab (J-PAL). https://www.povertyactionlab.org/policy-insight/microcredit-impacts-and-promising-innovations

Jui, T. D., & Rivas, P. (2024). Fairness issues, current approaches, and challenges in machine learning models. *International Journal of Machine Learning and Cybernetics*, *15*(8), 3095–3125. https://doi.org/10.1007/s13042-023-02083-2

Kabeer, N. (1999). Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment. *Development and Change*, *30*(3), 435–464. https://doi.org/10.1111/1467-7660.00125

Kanno-Youngs, Z., & Kliff, S. (2024, June 11). Biden Proposes Dropping Medical Debt From Credit Reports. *The New York Times*. https://www.nytimes.com/2024/06/11/us/politics/biden-medical-debt-credit-reports.html

Karlan, D., & Zinman, J. (2011). Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation. *Science*, *332*(6035), 1278–1284. https://doi.org/10.1126/science.1200138

Kaufman, M. R., Eschliman, E. L., & Karver, T. S. (2023). Differentiating sex and gender in health research to achieve gender equity. *Bulletin of the World Health Organization*, *101*(10), 666–671. https://doi.org/10.2471/BLT.22.289310

Kearns, J. (2023). *AI's Reverberations across Finance*. IMF. https://www.imf.org/en/Publications/fandd/issues/2023/12/AI-reverberations-across-finance-Kearns

Kelly, S., & Mirpourian, M. (2021). Algorithmic Bias, Financial Inclusion, and Gender. *Women's World Banking*.

Kim, M., & and Duvendack, M. (2024). Digital credit for all? An empirical analysis of mobile loans for financial inclusion in Kenya. *Information Technology for Development*, 1–18. https://doi.org/10.1080/02681102.2024.2402996

Kiviat, B. (2019). The Moral Limits of Predictive Practices: The Case of Credit-Based Insurance Scores. *American Sociological Review*, *84*(6), 1134–1158. https://doi.org/10.1177/0003122419884917

Kleinberg, J., Ludwig, J., Mullainathan, S., & Rambachan, A. (2018). Algorithmic Fairness. *AEA Papers and Proceedings*, *108*, 22–27. https://doi.org/10.1257/pandp.20181018

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2016). *Inherent Trade-Offs in the Fair Determination of Risk Scores* (arXiv:1609.05807). arXiv. https://doi.org/10.48550/arXiv.1609.05807

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, *117*(14), 7684–7689. https://doi.org/10.1073/pnas.1915768117

Koulish, R., & Evans, K. (2020). Injustice and the Disappearance of Discretionary Detention under Trump: Detaining Low Risk Immigrants without Bond. *ILCSS Working Paper*, *5*. https://scholarship.law.duke.edu/faculty_scholarship/4006

Kozodoi, N., Jacob, J., & Lessmann, S. (2022). Fairness in credit scoring: Assessment, implementation and profit implications. *European Journal of Operational Research*, *297*(3), 1083–1094. https://doi.org/10.1016/j.ejor.2021.06.023

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyńska, K. (2023).

What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine. *Diagnostics*, *13*(15), 2582. https://doi.org/10.3390/diagnostics13152582

Kuhn, T. S. (with Ralph Ellison Collection (Library of Congress)). (1970). *The structure of scientific revolutions* ([2d ed., enl). University of Chicago Press.

Lamsaf, A., Carrilho, R., Neves, J. C., & Proença, H. (2025). Causality, Machine Learning, and Feature Selection: A Survey. *Sensors*, *25*(8), Article 8. https://doi.org/10.3390/s25082373

Langenbucher, K. (2020). Responsible A.I.-based Credit Scoring – A Legal Framework. *European Business Law Review*, *31*(4). https://kluwerlawonline.com/api/Product/CitationPDFURL?file=Journals\EULR\EULR2020022.pdf

Langley, P., & Leyshon, A. (2022). Neo-colonial credit: FinTech platforms in Africa. *Journal of Cultural Economy*, *15*(4), 401–415. https://doi.org/10.1080/17530350.2022.2028652

Larrazabal, A. J., Nieto, N., Peterson, V., Milone, D. H., & Ferrante, E. (2020). Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proceedings of the National Academy of Sciences*, *117*(23), 12592–12594. https://doi.org/10.1073/pnas.1919012117

Larson, B. (2017). Gender as a Variable in Natural-Language Processing: Ethical Considerations. In D. Hovy, S. Spruit, M. Mitchell, E. M. Bender, M. Strube, & H. Wallach (Eds.), *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing* (pp. 1–11). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-1601

Lauer, J. (2017). *Creditworthy: A History of Consumer Surveillance and Financial Identity in America* (p. 368 Pages). Columbia University Press.

Lee, S., & Persson, P. (2024). *Theoretical Perspectives on Social Relationships in Informal Financing* (SSRN Scholarly Paper 4728615). Social Science Research Network. https://doi.org/10.2139/ssrn.4728615

Li, G. (2018). Gender-Related Differences in Credit Use and Credit Scores. *US Federal Reserve System*. https://www.federalreserve.gov/econres/notes/feds-notes/gender-related-differences-in-credit-use-and-credit-scores-20180622.html

Liu, L. T., Dean, S., Rolf, E., Simchowitz, M., & Hardt, M. (2018). Delayed Impact of Fair Machine Learning. *Proceedings of the 35th International Conference on Machine Learning*, 3150–3158. https://proceedings.mlr.press/v80/liu18c.html

Luo, C., Wu, D., & Wu, D. (2017). A deep learning approach for credit scoring using credit default swaps. *Engineering Applications of Artificial Intelligence*, *65*, 465–470. https://doi.org/10.1016/j.engappai.2016.12.002

Ma, S., Seidi, D., & McNulty, T. (2020). Challenges and practices of interviewing business elites. *Strategic Organization*, *19*(1). https://journals.sagepub.com/doi/10.1177/1476127020980969

MacKenzie, D., & Wajcman, J. (1999). *The Social Shaping of Technology*. Oxford University Press. https://www.research.ed.ac.uk/en/publications/the-social-shaping-of-technology-3

Marcuse, H. (1964). *One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society*. Routledge. https://doi.org/10.4324/9780203995211

Marda, V., & Narayan, S. (2021). On the importance of ethnographic methods in AI research. *Nature Machine Intelligence*, *3*(3), 187–189. https://doi.org/10.1038/s42256-021-00323-0

Marriage Rates by County—Prevalence of Monogamous and Polygamous Marriag. (2022). Kenya Data & Statistics. https://statskenya.co.ke/at-stats-kenya/about/marriage-rates-in-kenya-by-county/60/

Marron, D. (2015). Debt, consumption and freedom: Social scientific representations of consumer credit in Anglo-America. *History of the Human Sciences*, *28*(4), 25–43. https://doi.org/10.1177/0952695115599314

McGregor, S., Paeth, K., & Lam, K. (2022). *Indexing AI Risks with Incidents, Issues, and Variants* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2211.10384

Meager, R. (2022). Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature. *American Economic Review*, *112*(6), 1818–1847.

https://doi.org/10.1257/aer.20181811

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019). *A Survey on Bias and Fairness in Machine Learning* (arXiv:1908.09635). arXiv. http://arxiv.org/abs/1908.09635

Merriam-Webster. (2025, May 29). *Definition of fairness*. Merriam-Webster. https://www.merriam-webster.com/dictionary/fairness

Mhlanga, D. (2021). Financial Inclusion in Emerging Economies: The Application of Machine Learning and Artificial Intelligence in Credit Risk Assessment. *International Journal of Financial Studies*, *9*(3), Article 3. https://doi.org/10.3390/ijfs9030039

Miasato, A., & Reis, F. (2020). Artificial Intelligence as an Instrument of Discrimination in Workforce Recruitment | Request PDF. *Acta Universitatis Sapientiae Legal Studies*, *8*(2), 191–212. https://doi.org/10.47745/AUSLEG.2019.8.2.04

MICDE. (2025). *Kenya Artificial Intelligence Strategy 2025-2030*. Kenyan Ministry of Information, Communications and the Digital Economy (MICDE). https://www.ict.go.ke/

MicroSave. (2019, September 17). *Making digital credit truly responsible- Insights from Kenya*. MicroSave Consulting (MSC). https://www.microsave.net/2019/09/18/11310/

Miles, A. (2015, November 16). *From microfinance to financial inclusion: Reflections on 20 years*. CGAP. https://www.cgap.org/blog/microfinance-to-financial-inclusion-reflections-on-20-years, https://www.cgap.org/blog/microfinance-to-financial-inclusion-reflections-on-20-years

Mirpourian, M., Fu, J., & Kelly, S. (2022). Check Your Bias A Field Guide for Lenders. *Women's World Banking*.

Mobile Operating System Market Share Africa. (2024). StatCounter Global Stats. https://gs.statcounter.com/os-market-share/mobile/africa

Mobile Operating System Market Share Asia. (2024). StatCounter Global Stats. https://gs.statcounter.com/os-market-share/mobile/asia

Mohanty, C. T. (1984). Under Western Eyes: Feminist Scholarship and Colonial Discourses. *Boundary 2*, *12*(On Humanism and the University: The Discourse of Humanism), 338–358.

Morawczynski, O., & Pickens, M. (2009). *Poor People Using Mobile Financial Services: Observations on Customer Usage and Impact from M-PESA*. World Bank. https://openknowledge.worldbank.org/entities/publication/409abe95-c397-5ba6-8beb-9318ed90920e

Moreo, A., Esuli, A., & Sebastiani, F. (2019). *Word-Class Embeddings for Multiclass Text Classification* (arXiv:1911.11506). arXiv. https://doi.org/10.48550/arXiv.1911.11506

Morozov, E. (with Internet Archive). (2013). *To save everything, click here: The folly of technological solutionism*. PublicAffairs. http://archive.org/details/tosaveeverything0000moro

Moser, C. (2012). *Gender Planning and Development: Theory, Practice and Training*. Routledge. https://doi.org/10.4324/9780203411940

Mukhanova, A., Baitemirov, M., Amirov, A., Tassuov, B., Makhatova, V., Kaipova, A., Makhazhanova, U., & Ospanova, T. (2024). Forecasting creditworthiness in credit scoring using machine learning methods. *International Journal of Electrical and Computer Engineering (IJECE)*, *14*(5), Article 5. https://doi.org/10.11591/ijece.v14i5.pp5534-5542

Mulligan, D., Kroll, J. A., Kohli, N., & Wong, R. (2019). This Thing Called Fairness: Disciplinary Confusion Realizing a Value in Technology: Proceedings of the ACM on Human-Computer Interaction: Vol 3, No CSCW. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–36.

Mungiria, J. B., & Ondabu, I. T. (2019). Role of Credit Reference Bureau on Financial Intermediation: Evidence from the Commercial Banks in Kenya. *International Journal of Finance and Accounting*, *8*(3), 73–79.

Mutua, J. (2021, February 28). *MPs approve new law to regulate mobile loan rates*. Business Daily. https://www.businessdailyafrica.com/bd/corporate/companies/mps-approve-regulate-mobile-loan-rates-3307434

Nader, L. (1972). *Up the Anthropologist: Perspectives Gained From Studying Up*. https://eric.ed.gov/?id=ED065375

National Council For Population and Development. (2017). *Policy Brief: Youth Bulge in Kenya*. National Council For Population and Development. https://ncpd.go.ke/policy-briefs/

Newton, G., Kirby, E., Hofstätter, L., Judd-Lam, S., Smith, L., Churchill, B., Strnadová, I., & Newman, C. E. (2024). 'Is There Anything Else You'd Like to Tell Us About Your Experience?' Orientations Towards Listening to Open-Ended Survey Responses. *Sociological Research Online*, 13607804241287628. https://doi.org/10.1177/13607804241287628

NITI Aayog. (2018). *National Strategy for Artificial Intelligence*. National Institution for Transforming India (NITI). https://indiaai.gov.in/

NITI Aayog. (2021). *Responsible AI #AIFORALL*. National Institution for Transforming India (NITI). https://www.niti.gov.in/

Nkrumah, K. (1965). *Neo-Colonialism, the Last Stage of imperialism*. International Publishers Co. https://www.marxists.org/subject/africa/nkrumah/neo-colonialism/

Noble, D., & Stalder, F. (1998). The religion of technology: The divinity of man & the spirit of invention. *Canadian Journal of Communication*, *23*(4), 557–560.

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press. https://doi.org/10.2307/j.ctt1pwt9w5

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science (New York, N.Y.)*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

O'Connor, S., & Liu, H. (2024). Gender bias perpetuation and mitigation in AI technologies: Challenges and opportunities. *AI & SOCIETY*, *39*(4), 2045–2057. https://doi.org/10.1007/s00146-023-01675-4

O'Donnell, A., & and Sweetman, C. (2018). Introduction: Gender, development and ICTs. *Gender & Development*, *26*(2), 217–229. https://doi.org/10.1080/13552074.2018.1489952

OECD. (2024). *FinTech lending in Sub-Saharan Africa*. OECD.

OECD, & ILO. (2019). *Addressing the gender dimension of informality* (pp. 131–154). OECD.

https://doi.org/10.1787/cfd32100-en

Oldenziel, R. (1999). *Making Technology Masculine: Men, Women, and Modern Machines in America, 1870-1945*. Amsterdam University Press. https://www.jstor.org/stable/j.ctt46mtdk

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group. https://www.amazon.com/Weapons-Math-Destruction-Increases-Inequality/dp/0553418815

Oosterlaken, I. (2009). Design for Development: A Capability Approach. *Design Issues*, *25*(4), 91–102. https://doi.org/10.1162/desi.2009.25.4.91

Oosterlaken, I. (2011). Inserting Technology in the Relational Ontology of Sen's Capability Approach. *Journal of Human Development and Capabilities*, *12*(3), 425–432. https://doi.org/10.1080/19452829.2011.576661

O'Reilly, M., Shemluck, N., & Cooper, L. (2025, April 16). *How the Bay Area became an epicenter of fintech innovation*. San Francisco Business Times. https://www.bizjournals.com/sanfrancisco/news/2025/04/16/bay-area-epicenter-fintech-innovation.html

Orlikowski, W. J. (1992). The Duality of Technology: Rethinking the Concept of Technology in Organizations. *Organization Science*, *3*(3), 398–427.

Óskarsdóttir, M., Bravo, C., Sarraute, C., Baesens, B., & Vanthienen, J. (2020). *Credit Scoring for Good: Enhancing Financial Inclusion with Smartphone-Based Microlending* (arXiv:2001.10994). arXiv. https://doi.org/10.48550/arXiv.2001.10994

Palmié, M., Wincent, J., Parida, V., & Caglar, U. (2020). The evolution of the financial technology ecosystem: An introduction and agenda for future research on disruptive innovations in ecosystems. *Technological Forecasting and Social Change*, *151*, 119779. https://doi.org/10.1016/j.techfore.2019.119779

Parpart, J., Connelly, P., & Barriteau, E. (2000). *Theoretical Perspectives on Gender and Development*. IDRC - International Development Research Centre.

https://idrc-crdi.ca/en/books/theoretical-perspectives-gender-and-development

Pasquale, F. (2019, November 25). *The Second Wave of Algorithmic Accountability*. LPE Project. https://lpeproject.org/blog/the-second-wave-of-algorithmic-accountability/

Perez, S. (2023, November 9). *Google Play tightens up rules for Android app developers to require testing, increased app review*. TechCrunch. https://techcrunch.com/2023/11/09/google-play-tightens-up-rules-for-android-app-developers-to-require-testing-increased-app-review/

Peria, M., Soledad, M., & Singh, S. (2014). *The impact of credit information sharing reforms on firm financing?* [Text/HTML]. World Bank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/en/437081468340849287

Pesqué-Cela, V., Tian, L., Luo, D., Tobin, D., & Kling, G. (2021). Defining and measuring financial inclusion: A systematic review and confirmatory factor analysis. *Journal of International Development*, *33*(2), 316–341. https://doi.org/10.1002/jid.3524

Pessach, D., & Shmueli, E. (2022). A Review on Fairness in Machine Learning. *ACM Comput. Surv.*, *55*(3), 51:1-51:44. https://doi.org/10.1145/3494672

Philip, S. (2018). Youth and ICTs in a ‘new’ India: Exploring changing gendered online relationships among young urban men and women. *Gender & Development*, *26*(2), 313–324. https://doi.org/10.1080/13552074.2018.1473231

Pomeranz, D. (2014). *The Promise of Microfinance and Women's Empowerment: What Does the Evidence Say?* [Discussion Paper]. https://www.hbs.edu/faculty/Pages/item.aspx?num=47159

Postelnicu, L., & Hermes, N. (2018). Microfinance Performance and Social Capital: A Cross-Country Analysis. *Journal of Business Ethics*, *153*(2), 427–445. https://doi.org/10.1007/s10551-016-3326-0

Prasad, A. (2014). *Imperial Technoscience*. MIT Press. https://mitpress.mit.edu/9780262026956/imperial-technoscience/

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731. https://doi.org/10.3758/BF03206553

Procaccini, F. (2015). Stemming the Rising Risk of Credit Inequality. *Harvard Law and Policy Review*, *9*, 43.

Qureshi, S. (2020). Why Data Matters for Development? Exploring Data Justice, Micro-Entrepreneurship, Mobile Money and Financial Inclusion. *Information Technology for Development*, *26*(2), 201–213. https://doi.org/10.1080/02681102.2020.1736820

Reddy, K., Wallace, D., & Wellalage, N. H. (2024). The impact of financial literacy on financial inclusion. *Australian Journal of Management*, 03128962241270809. https://doi.org/10.1177/03128962241270809

Reinharz, S. (1992). *Feminist Methods in Social Research*. Oxford University Press.

Rennie, S., Buchbinder, M., Juengst, E., Brinkley-Rubinstein, L., Blue, C., & Rosen, D. L. (2020). Scraping the Web for Public Health Gains: Ethical Considerations from a 'Big Data' Research Project on HIV and Incarceration. *Public Health Ethics*, *13*(1), 111–121. https://doi.org/10.1093/phe/phaa006

Reserve Bank of India. (2014). *Report of the Committee to Recommend Data Format for Furnishing of Credit Information to Credit Information Companies*. Reserve Bank of India. https://rbi.org.in/scripts/PublicationReportDetails.aspx?UrlPage=&ID=763#215

Rice, L., & Swesnik, D. (2012). Discriminatory Effects of Credit Scoring on Communities of Color. *National Fair Housing Alliance*.

Robinson, D. (2022). *Voices in the Code*. Russell Safe Foundation. https://www.russellsage.org/publications/voices-code

Robinson, J., Park, D. S., & Blumenstock, J. E. (2023). *The Impact of Digital Credit in Developing Economies: A Review of Recent Evidence* (SSRN Scholarly Paper 4540063). Social Science Research Network. https://doi.org/10.2139/ssrn.4540063

Rothchild, J. (2014). Gender Bias. In *The Blackwell Encyclopedia of Sociology*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405165518.wbeosg011.pub2

Rowntree, O., Bahia, K., & Butler, C. (2020). *The Mobile Gender Gap Report 2020*. https://data.gsmaintelligence.com/research/research/research-2020/the-mobile-gender-gap-report-2020

Sadok, H., Sakka, F., & El Maknouzi, M. E. H. (2022). Artificial intelligence and bank credit analysis: A review. *Cogent Economics & Finance*, *10*(1), 2023262. https://doi.org/10.1080/23322039.2021.2023262

Sadowski, J. (2025). Machine's Eye View: Postmodern Data Science and the Politics of Ground Truth. *Science, Technology, & Human Values*, 01622439251331138. https://doi.org/10.1177/01622439251331138

Said, E. (1978). *Orientalism*. Penguin Random House. https://www.penguinrandomhouse.com/books/159783/orientalism-by-edward-w-said/

Salamina, L. M., Chhabra, P., Sankaranarayan, S., & Masunda, C. (2019). *Disruptive Technologies in the Credit Information Sharing Industry: Developments and Implications* [Text/HTML]. World Bank. https://documents.banquemondiale.org/fr/publication/documents-reports/documentdetail/en/587611557814694439

Santamaría, L., & Mihaljević, H. (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ. Computer Science*, *4*, e156. https://doi.org/10.7717/peerj-cs.156

Schech, S. (2002). Wired for change: The links between ICTs and development discourses. *Journal of International Development*, *14*(1), 13–23.

Scheffler, P. (2025, February 11). Machine Learning in Fintech: Use Cases & Implementation Steps. *Neontri*. https://neontri.com/blog/machine-learning-fintech/

Schicks, J. (2013). The Definition and Causes of Microfinance Over-Indebtedness: A Customer Protection Point of View. *Oxford Development Studies*, *41*(sup1), S95–S116. https://doi.org/10.1080/13600818.2013.778237

Schreiber, D. (2019). *How AI Can Vanquish Bias*. Insurance Thought Leadership. https://www.insurancethoughtleadership.com/ai-machine-learning/how-ai-can-vanquish-bias

Schwemmer, C., Knight, C., Bello-Pardo, E. D., Oklobdzija, S., Schoonvelde, M., & Lockhart, J. W. (2020). Diagnosing Gender Bias in Image Recognition Systems. *Socius*, *6*, 2378023120967171. https://doi.org/10.1177/2378023120967171

Scientific Consensus on AI Bias. (2025). Scientific Consensus on AI Bias. https://www.aibiasconsensus.org

Sen, A. (1999). *Development as Freedom*. Anchor. https://www.amazon.com/Development-as-Freedom-Amartya-Sen/dp/0385720270

Sen, A. (2010). The Mobile and the World. *Information Technologies & International Development*, *6*(SE), Article SE.

Shahriar, A. Z. M., Unda, L. A., & Alam, Q. (2020). Gender differences in the repayment of microcredit: The mediating role of trustworthiness. *Journal of Banking & Finance*, *110*, 105685. https://doi.org/10.1016/j.jbankfin.2019.105685

Shipton, L., & Vitale, L. (2024). Artificial intelligence and the politics of avoidance in global health. *Social Science & Medicine*, *359*, 117274. https://doi.org/10.1016/j.socscimed.2024.117274

Showkat, M., Nagina, R., Baba, M. A., & Yahya, A. T. (2025). The impact of financial literacy on women's economic empowerment: Exploring the mediating role of digital financial services. *Cogent Economics & Finance*, *13*(1), 2440444. https://doi.org/10.1080/23322039.2024.2440444

Singh, J. (2022, August 26). Predatory loan apps in India driving some users to suicide. *TechCrunch*. https://techcrunch.com/2022/08/26/loan-apps-abuse-harassment-suicide-indian-users-google-apple-india/

Smart, A., & Kasirzadeh, A. (2024). *Beyond Model Interpretability: Socio-Structural Explanations in Machine Learning* (arXiv:2409.03632). arXiv. https://doi.org/10.48550/arXiv.2409.03632

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology*, *13*, 290–312. https://doi.org/10.2307/270723

Sohn, S. Y., & Kim, J. W. (2012). Decision tree-based technology credit scoring for start-up firms: Korean case. *Expert Systems with Applications*, *39*(4), 4007–4012. https://doi.org/10.1016/j.eswa.2011.09.075

S&P Global. (2024). *Fintech funding falls 42% to $35B in 2023, but downturn may be nearing end*. https://www.spglobal.com/marketintelligence/en/news-insights/research/fintech-funding-falls-42-to-35b-in-2023-but-downturn-may-be-nearing-end

Statista. (2019). *Kenya: Population by highest level of education*. Statista. https://www.statista.com/statistics/1237796/distribution-of-population-in-kenya-by-highest-level-of-education-completed/

Subramaniam, B., Foster, L., Harding, S., Roy, D., & TallBear, K. (2017). Feminism, Postcolonialism, and Technoscience. In *Handbook of Science and Technology Studies*. MIT Press. https://genderstudies.indiana.edu/research/books/book-chapters/feminism-postcolonialism-and-technoscience.html

Suri, T., Bharadwaj, P., & Jack, W. (2021). Fintech and household resilience to shocks: Evidence from digital loans in Kenya. *Journal of Development Economics*, *153*, 102697. https://doi.org/10.1016/j.jdeveco.2021.102697

Tay, L., Woo, S. E., Hickman, L., Booth, B. M., & D'Mello, S. (2022). A Conceptual Framework for Investigating and Mitigating Machine-Learning Measurement Bias (MLMB) in Psychological Assessment. *Advances in Methods and Practices in Psychological Science*, *5*(1), 25152459211061337. https://doi.org/10.1177/25152459211061337

The White House. (2025, July 23). *Preventing Woke AI in the Federal Government*. The White House. https://www.whitehouse.gov/presidential-actions/2025/07/preventing-woke-ai-in-the-federal-government/

Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, B., Belgrave, D. C. M., Ezer, D., Haert, F. C. van der, Mugisha, F., Abila, G., Arai, H., Almiraat, H., Proskurnia, J., Snyder, K., Otake-Matsuura, M., Othman, M., Glasmachers, T., Wever, W. de, … Clopath, C.

(2020). AI for social good: Unlocking the opportunity for positive impact. *Nature Communications*, *11*(1), 2468. https://doi.org/10.1038/s41467-020-15871-z

Tongia, R., Subrahmanian, E., & Arunachalam, V. (2005). *Information and Communications Technology for Sustainable Development: Defining a Global Research Agenda*. Allied Publishers. https://www.nist.gov/publications/information-and-communications-technology-sustainable-development-defining-global

TransUnion Africa. (2022). *TransUnion Kenya, Metropol CRB and Creditinfo CRB Kenya Highlight Opportunities Made Possible by Credit Scoring*. TransUnion Africa. https://newsroom.transunionafrica.com/transunion-kenya-metropol-crb-and-creditinfo-crb-kenya-highlight-opportunities-made-possible-by-credit-scoring/

Ukibe, B. (2024, February 22). Branch Pro: Meet the First Neobank in Kenya. *Silicon Africa*. https://siliconafrica.org/branch-pro/

UN Women. (2023). *Progress on the Sustainable Development Goals: The gender snapshot 2023*. UN Women. https://www.unwomen.org/en/digital-library/publications/2023/09/progress-on-the-sustainable-development-goals-the-gender-snapshot-2023

UNESCO. (2020, June 22). *Gender equity definition*. UNESCO. https://uis.unesco.org/en/glossary-term/gender-equity

US Department of Justice. (2023). *Living arrangements of children by race/ethnicity, 1970-2023*. Office of Juvenile Justice and Delincuency Prevention. https://www.ojjdp.gov/ojstatbb//population/qa01202.asp?qaDate=2023

Vasiliev, Y. (2020). *Natural Language Processing with Python and spaCy*. No Starch Press. https://nostarch.com/NLPPython

Veale, M., Van Kleek, M., & Binns, R. (2018). *Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making* (SSRN Scholarly Paper 3175424). Social Science Research Network. https://papers.ssrn.com/abstract=3175424

Vuković, D. B., Dekpo-Adza, S., & Matović, S. (2025). AI integration in financial services: A systematic review of trends and regulatory challenges. *Humanities and Social Sciences Communications*, *12*(1), 1–29. https://doi.org/10.1057/s41599-025-04850-8

Wajcman, J. (2002). Addressing Technological Change: The Challenge to Social Theory. *Current Sociology*, *50*(3), 347–363. https://doi.org/10.1177/0011392102050003004

Wajcman, J. (2006). *TechnoFeminism* (repr). Polity.

Wang, K., Li, M., Cheng, J., Zhou, X., & Li, G. (2022). Research on personal credit risk evaluation based on XGBoost. *Procedia Computer Science*, *199*, 1128–1135. https://doi.org/10.1016/j.procs.2022.01.143

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

West, D. (2018). *What is artificial intelligence?* Brookings. https://www.brookings.edu/articles/what-is-artificial-intelligence/

West, S. M. (2019, April 1). Discriminating Systems: Gender, Race, and Power in AI - Report. *AI Now Institute*. https://ainowinstitute.org/publication/discriminating-systems-gender-race-and-power-in-ai-2

Williams, B. A., Brooks, C. F., & Shmargad, Y. (2018). How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications. *Journal of Information Policy*, *8*, 78–115. https://doi.org/10.5325/jinfopoli.8.2018.0078

World Bank. (n.d.). *Financial Inclusion* [Text/HTML]. World Bank. Retrieved June 6, 2025, from https://www.worldbank.org/en/topic/financialinclusion/overview

World Bank. (2013). *Global financial development report 2014: Financial inclusion*. World Bank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/225251468330270218/global-financial-development-report-2014-financial-inclusion

World Bank. (2021). *The Global Findex Database 2021* [Text/HTML]. World Bank.

https://www.worldbank.org/en/publication/globalfindex/Report

World Bank. (2024). *Financial Inclusion in Sub-Saharan Africa* [Text/HTML]. World Bank. https://www.worldbank.org/en/publication/globalfindex/brief/financial-inclusion-in-sub-saharan-africa-overview

World Bank Group. (2023). *World Bank Open Data*. World Bank Open Data. https://data.worldbank.org

World Economic Forum. (2023). *Global Gender Gap Report 2023*. World Economic Forum. https://www.weforum.org/publications/global-gender-gap-report-2023/in-full/

World Economic Forum. (2024). *Global Gender Gap Report 2024*. World Economic Forum. https://www.weforum.org/publications/global-gender-gap-report-2024/

Yang, Y., Zhang, H., Gichoya, J. W., Katabi, D., & Ghassemi, M. (2024). The limits of fair medical imaging AI in real-world generalization. *Nature Medicine*, 1–11. https://doi.org/10.1038/s41591-024-03113-4

Young, E., Wajcman, J., & Sprejer, L. (2021). *Where are the women? Mapping the gender job gap in AI*. The Alan Turing Institute. https://www.turing.ac.uk/news/publications/report-where-are-women-mapping-gender-job-gap-ai

Young, R. J. C. (2020). Postcolonialism: A Very Short Introduction. In R. J. C. Young (Ed.), *Postcolonialism: A Very Short Introduction* (2nd ed., p. 0). Oxford University Press. https://doi.org/10.1093/actrade/9780198856832.003.0001

Yunus, M. (2004). Grameen Bank, Microcredit and Millennium Development Goals. *Economic and Political Weekly*, *39*(36), 4–10. https://doi.org/10.2307/4415509

Zaidan, E., & Ibrahim, I. A. (2024). AI Governance in a Complex and Rapidly Changing Regulatory Landscape: A Global Perspective. *Humanities and Social Sciences Communications*, *11*(1), 1–18. https://doi.org/10.1057/s41599-024-03560-x

Zou, Y., Xia, M., & Lan, X. (2025). Interpretable credit scoring based on an additive extreme gradient boosting. *Chaos, Solitons & Fractals*, *194*, 116216. https://doi.org/10.1016/j.chaos.2025.116216

APPENDIX

Appendix 1. Interview Questions & Topic Guide

A. Interview questions for corporate leaders

Background & conceptualization

1. Tell me about how the company was started: What was the problem the founder was trying to solve and why did they want to solve it?
    a. Probe: Why did the team decide to use machine learning (ML) for lending?
    b. Probe: How does the company make money / what is the business model?
2. What brought you to the organization and your role?

Tool Design

3. What is the purpose of the ML model and what is it trying to predict?
4. How does the ML model work? What data does the model collect from people to assess one's credit and how is the data collected?
    a. Probe: Who are the target users, and why target them?
5. Who was / is involved in the design of the ML model?
    a. Probe: Demographic makeup of the design team (gender, other identities), where based, and expertise
    b. Probe: Were any stakeholders outside of the organization (e.g., community members) involved in the design of the model? If so, how and why were they included?

Tool Management & Use

6. Can you tell me more about the management team?
    a. Probe: Demographic makeup (gender, other identities), where based, and expertise
7. To whom are the credit scores given in order to facilitate loans? How do the loan providers determine terms (e.g., loan size, interest rates, repayment period)?
8. Is the ML model adapted for different countries and/or contexts? If Y, please share more about this.
    a. Probe: What challenges have been encountered in expanding to different places and how, if at all, has the organization changed tack?
9. How does the team work to ensure the model is operating fairly?
    a. Probe: What does this look like? How about in regards to gender?
10. [Potential to skip] Does the organization have values or principles it follows, and if so, what are these? (*For example: "do no harm"*)
    a. Probe: How do those manifest (or not) in the design and management of the ML tool?

User Demographics

11. Can you share the breakdown by country of…
    a. Users / applicants (# M/W/nonbinary)
    b. [if applicable] borrowers (# M/W/nonbinary)
    c. [if applicable] average loan size for M/W/nonbinary

d. Probe: Can you share a breakdown including other identity categories if tracked? (e.g., socio-economic status, caste)
e. Note: If this has already been answered by an interview with a data scientist at the company, then it will be skipped

12. Are there any patterns in usage of your lending products by gender or other demographics? *For example, do more men tend to access the app and face less challenges in navigating the app.*
    a. Probe: Why do you think that is?

Lessons & Impacts

13. Does the organization track impacts of loans on borrowers, and if so, what are impacts of these loans on borrowers? Do impacts differ by gender, and if so, how?
    a. Probe: If yes, why do you think impacts differ by gender?
14. What lessons have you learned about the advantages and limitations of employing ML in pursuit of financial inclusion in the countries (or relevant regions) that you operate?
    a. Probe: In what ways is the organization enhancing financial inclusion, and for whom?
    b. Probe: What challenges does the organization face in efforts to enhance financial inclusion, particularly related to gender?
    c. Probe: In thinking about your or the organization's goals for the future, have they changed since you started, and what hopes do you have?
15. Before we finish, I'd like to make sure we've covered the important pieces important in thinking about impacts of machine learning alternative lending tools on gender equity. Any experiences or other information you would like to add?

Follow-up

16. Is there a data scientist at the organization I could interview? Is there anyone else you might recommend me to interview?
17. [For potential case study companies]
    a. Are there any reports that might be helpful for me to read (internal or external)?
    b. Would it be possible to see the data on who is applying for loans and credit scores? (Can sign NDA and would be anonymous)
    c. [For those of interest] Would it be possible to conduct a survey with your users to explore more around how they are accessing and using the app, as well as any gender differences?

B. Interview questions for data scientists

Background & conceptualization

1. What brought you to the organization and your role?
2. Tell me about the motivations in using machine learning (ML): Why did the team decide to use ML for lending and how has the approach evolved over time, if at all?

Design

3. What is the purpose of the ML model and what is it trying to predict?
    a. Example: The purpose is to provide a credit score; it tries to predict willingness to repay and trust
4. How does the ML model work?
    a. Probe: Does it use supervised, unsupervised and/or reinforcement learning?
    b. Probe: What data was the model trained on? What about evaluated on?

c. Probe: What variables or proxies are being used to assess creditworthiness?
d. Probe: What features are found to be most important in the model?

5. What data does the model collect from people to assess one's credit and how is the data collected?
6. Who was / is involved in the design of the ML model?
   a. Probe: Demographic makeup of the design team (gender, other identities), where based, and expertise
   b. Probe: Were any stakeholders outside of the organization (e.g., community members) involved in the design of the model?
7. How has gender been considered, or not, in the design and development of the tool?

Use & management

8. To whom are the credit scores given in order to facilitate loans? How do the loan providers determine loan size and interest rates?
9. Is the ML tool adapted for different countries and/or contexts? If Y, please share more about this.
10. How does the team work to ensure and assess whether the model is operating fairly?
    a. Probe: Does the organization audit or regularly assess the ML tool? If Y, please share more about the auditing process and any findings related to gender. *(E.g., in credit scores, in credit offers, in loan sizes / interest rates (when controlling for relevant variables))*
    b. Probe: When a credit score is provided from the ML tool, is it possible to identify how the score was generated?
11. Have you observed any instances of bias in your tool? If so, what bias was found and how did the team work to address it or not? *(E.g., bias in data (representation, societal bias), bias in algorithms (proxies used), bias in how used (lenders interpreting results)*
    a. Probe: What about bias specific to gender?
    b. Probe: Do you think diversity on your team is helpful to mitigate gender bias? If so, in what ways?
12. [Potential to skip] What type of information is shown or available to users about their credit assessment and how it impacts their offerings?

User Demographics

13. Can you share the breakdown of…
    a. Users / applicants (# M/W/nonbinary); by region
    b. [if applicable] borrowers (# M/W/nonbinary)
    c. [if applicable] average loan size for M/W/nonbinary
    d. Probe: Please share breakdown including other identity categories if tracked (e.g., socio-economic status, caste)
14. Are there any patterns in usage of your lending products by gender or other demographics? *For example, do more men tend to access the app and face less challenges in navigating the app.*

Lessons & Impacts

15. Does the organization track impacts of loans on borrowers, and if so, what are impacts of these loans on borrowers? How do impacts differ by gender?
    a. Probe: How do you assess whether the ML-based lending products are enhancing financial inclusion or not?
    b. Probe: Do you consider demographics in assessing financial inclusion (e.g., gender, caste)

16. What lessons have you learned about the advantages and limitations of employing ML in pursuit of financial inclusion in the countries (or relevant regions) that you operate?
    a. Probe: In thinking about your or the organization's goals for the future, have they changed since you started, and what hopes do you have?
17. Before we finish, I'd like to make sure we've covered the important pieces important in thinking about impacts of machine learning alternative lending tools on gender equity. Any experiences or other information you would like to add?

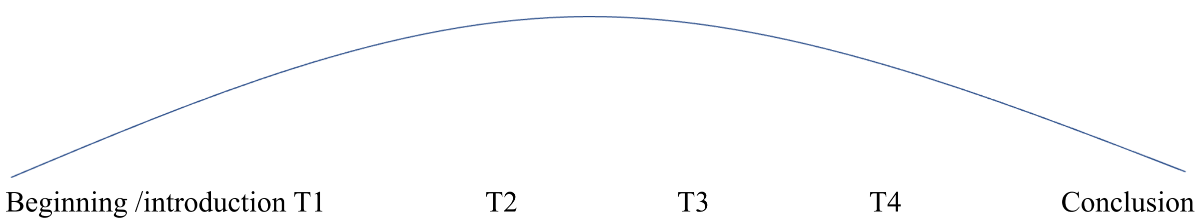
Follow-up

18. Is there anyone else you might recommend me to interview?

Topic Guide

While questions vary slightly based on the two audiences, the general topic trajectory remains the same.

T1: Background of organization & conceptualization of ML tool	● Fintech background & founding story ● Demographic background of organization & individual ● Purpose of using an ML tool for alternative lending
T2: Design of ML Tool	● Purpose of the ML tool and target users ● How the ML tool functions (e.g., approach, data, variables/proxies, fairness metrics)
T3: Use & management of ML tool	● How the model is used and by whom for loan facilitation ● Adapting the tool (or not) to different contexts ● Regular assessments or audits
T4: Lessons & impacts tracked by organization	● Reach of tool and its impacts ● Lessons learned



Appendix 2. Survey Questionnaire

Part 0. Survey overview & consent

An examination of machine learning-based alternative lending on financial inclusion & gender equity | CUREC Approval Reference: C1A_23_079

General Information
The aim of this research is to investigate whether and how lending apps that use artificial intelligence impact financial inclusion and gender equity in low and middle income countries. We appreciate your interest in participating in this survey.

You have been invited to participate as you self-identified meeting the conditions of the survey, which include (1) residing in Kenya; (2) being over 18 years old; and (3) having used the app, Tala. Please read through this information before agreeing to participate (if you wish to) by ticking the 'yes' box below.

You may ask any questions before deciding to take part by contacting the researcher (details below). The Principal Researcher is Genevieve Smith, who is attached to the Department of International Development at the University of Oxford. This research is being completed under the supervision of Professors Xiaolan Fu and Masooda Bano.

Participants will be asked to complete a short survey. This should take about 15 minutes. Upon completion of the survey, remuneration (KSh 400) will be provided.

No background knowledge is required. The data will help inform how people access and use the app, Tala, as well as any benefits and challenges faced. The data will be used by the Principal Researcher as part of the study to assess whether and how machine learning-based alternative lending tools impact financial inclusion and gender equity. No third parties will be given access to the data.

Do I have to take part?
No. Please note that participation is voluntary. If you do decide to take part, you may withdraw at any point for any reason before submitting your answers by pressing the 'Exit' button/ closing the browser. However, we are only able to compensate participants who complete the survey. After completing the survey, you are also able to withdraw from the study or withdraw personal information from the study through contacting the researcher with a withdraw request until data is anonymised.

How will my data be used?
The data we will collect that could identify you will be your gender and age. We will take all reasonable measures to ensure that data remain confidential. The responses you provide will be stored in a password-protected electronic file on University of Oxford secure servers and may be used in a thesis paper, academic publications, and conference presentations. At the end of the research project, the data will be deposited in Oxford's institutional data archive, ORA-Data.

Name and phone number will be collected at the end of the survey purely for payment purposes. Payment will be provided to your MPesa wallet via a money transfer service. Once payment has been processed, names and phone numbers will be destroyed.

Who will have access to my data?
The University of Oxford is the data controller with respect to your personal data and, as such, will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.
The results will be written up for a DPhil (PhD) degree.

Who has reviewed this research?
This research has been reviewed by, and received ethics clearance through, a subcommittee of the University of Oxford Central University Research Ethics Committee C1A_23_079.

Who do I contact if I have a concern or I wish to complain?
If you have a concern about any aspect of this research, please speak to Genevieve Smith (genevieve.smith@kellogg.ox.ac.uk) or their supervisor Xiaolan Fu (xiaolan.fu@qeh.ox.ac.uk), and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:
Prof. Loren Landau, The Chair, Department of International Development Research Ethics Committee; Email: loren.landau@qeh.ox.ac.uk; Address: 3 Mansfield Rd, Oxford OX1 3TB, UK

☐ I certify that I reside in Kenya and have used the app, Tala

Please note that you may only participate in this survey if you are 18 years of age or over.
☐ I certify that I am 18 years of age or over

If you have read the information above and agree to participate with the understanding that the data (gender, age) you submit will be processed accordingly, please tick the box below to start.
☐ Yes, I agree to take part

Part 1. Introductory self identification

- Please select your age range:
  - 18-24 years old
  - 25-34 years old
  - 35-44 years old
  - 45-54 years old
  - 55 years old or above
- Please select your gender:
  - Female
  - Male
  - Non-binary

- What city/town do you live in? [*open-ended*]
- Is the place you live now a large city (like Nairobi), a town (like Kikuyu), or a rural area such as a farming area?
  - Large city
  - Town
  - Rural area
- What best describes your main source of income?
  - Farming or agriculture (crops or livestock)
  - Employed
  - Casual worker
  - Self-employed / running own business
  - Receiving support from family/friends/spouse
  - Sub-letting of land or house
  - Receiving assistance from an aid agency/ NGO/ the government
  - No income
    - Unemployed
    - Student
  - Other:
- What best describes your education level?
  - No education
  - Some primary education
  - Completed primary education
  - Some secondary education
  - Completed secondary education
  - Vocational/Technical training (e.g. dressmaking, computer)
  - University degree
  - Other:
- What is your marital status?
  - Married
  - Widowed
  - Divorced
  - Unmarried / single
  - Other
- Have you done any of the following? Select all that apply
  - Received finance or a loan from an informal source or community group (e.g., through a community savings groups or Village Savings & Loan Association)
  - Received finance or a loan from a money lender
  - Received finance or a loan from a formal bank or neobank
  - Used a digital service for payments and transactions (e.g., MPesa)
  - Opened a savings account with a bank or neobank
  - Taken out formal insurance
  - None of the above
- Is Tala the only digital loan app that you have used? [Y/N]
- [If N] How many other digital loan apps have you tried or used?

Part 2. App usage

1. When did you download the Tala app (month, year)? [*open-ended*]
2. How many times have you used the app to get a loan?
    a. 0 - 1 – This is my first loan
    b. 2-5
    c. 6-9
    d. 10+
3. In regards to your latest loan from Tala, what was the size of the loan?
4. In regards to your latest loan from Tala, for what *purpose* did you take it out? (select all that apply)
    a. Purchase airtime
    b. Day to day living needs (food / transportation / rent / electricity bills or similar)
    c. Personal enjoyment (shopping, hobbies, gifts, or leisure activities)
    d. Travel or vacation expenses
    e. For medical needs including medical emergency for self, family, or friends
    f. For other emergencies such as home repairs, car repairs
    g. For education or school fees or supplies
    h. Start a business or side hustle
    i. Buy inventory/stock, supplies or other needs of existing business
    j. Farming supplies or resources
    k. New goods and appliances (phone, laptop, home appliance)
    l. Home improvement
    m. To pay off another debt, loan, or money owed
    n. Just to try it out
    o. Other:
5. Do you have other options to access a loan for the purpose(s) you needed?
    a. Yes
    b. No
    c. I don't know
6. In regards to your latest loan from Tala, how did you *actually* spend the money? (select all that apply)
    a. Purchase airtime
    b. Day to day living needs (food / transportation / rent / electricity bills or similar)
    c. Personal enjoyment (shopping, hobbies, gifts, or leisure activities)
    d. Travel or vacation expenses
    e. For medical needs including medical emergency for self, family, or friends
    f. For other emergencies such as home repairs, car repairs
    g. For education or school fees or supplies
    h. Start a business or side hustle
    i. Buy inventory/stock, supplies or other needs of existing business
    j. Farming supplies or resources
    k. New goods and appliances (phone, laptop, home appliance)
    l. Home improvement

    m. To pay off another debt, loan, or money owed
    n. Just to try it out
    o. Other:

Part 3. Impacts

7. Please rate how much you agree with the following statements. In thinking about the latest loan I took out from Tala, this loan and access to the Tala app has impacted… (1 = very negatively, reduced significantly; 2 = slightly negative, reduced somewhat; 3 = no change; 4 = slightly positively, increased somewhat; 5 = very positively, increased significantly)
    a. My self confidence
    b. How hopeful I am about the future
    c. My sense of control over how I spend money
    d. My ability to make decisions about how I spend money
    e. My ability to meet basic needs
    f. My financial well-being
    g. My involvement in household decisions for large purchases
8. If you took out the loan for a business reason, how did having the loan impact your business? (Skip if took out loan for personal use) (*open ended*)
9. Are there any other impacts you've experienced from using the Tala app and receiving a loan?

Part 4. Experience of using the app

10. What best describes the phone that you downloaded the Tala app on?
    a. My personal phone (I own it and am the sole user)
    b. A family phone (my family owns it and I use it along with other family members)
    c. Another person's phone
    d. Other
11. On a scale of 1 to 5 (1 = very easy, 5 = very challenging), how challenging was it for you to *download and use* the app?
12. Did you need any help from others in *downloading and using* the app initially?
    a. Yes, lots of help
    b. Yes, some help
    c. No
    d. I don't know
13. Have you experienced any of the following in using the Tala app? Check all that apply.
    a. The app wouldn't open or had glitches that made it difficult to use
    b. I didn't receive a credit / loan offer
    c. I received a credit / loan offer, but never received the loan monies in my bank account
    d. I didn't receive reminders for my loan payments
    e. Poor customer service in responding to my queries or asks for assistance
    f. None of the above
    g. Other:
14. On a scale of 1 to 5 (1 = Completely disagree; 2 = Somewhat disagree; 3 = Neither agree nor disagree; 4 = Somewhat agree; 5 = Completely agree), to what extent do you agree with the following…

a. I understood the loan approval process.
b. I understood all of the terms and conditions of Tala's digital loan, including payments and penalties.
c. I had **no** trouble understanding how to make loan payments.
d. I had **no** trouble affording your loan and interest payments?
e. I understand how the Tala app uses my data.
f. I trust Tala with my personal information.

15. Please explain your answers:
16. Were you ever late on payments for a loan secured through the app?
    a. Yes
    b. No
    c. I don't know

    [If Y] Why was the payment late?

    d. I forgot I took out a loan
    e. I forgot the due date
    f. I didn't have enough money
    g. It's too difficult/inconvenient to make the payment
    h. I was traveling
    i. Other:
17. Did you ever default on a loan secured through the app?
    a. Yes
    b. No
    c. I don't know
18. [If Y] Why did you default? (select all that apply)
    a. I forgot to make payments
    b. I had an unexpected expense
    c. Job or income loss
    d. The payment was more than expected
    e. Fees and interest became unaffordable
    f. Other:
19. Would you like to share any other information regarding your experience of the app? [*open ended, optional*]

Thank you for completing this survey!

Appendix 3. Variables for quantitative analysis

Appendix 3, Table 1. Independent and dependent variables

Variable	Variable type	Category	About
Female	Independent	Gender	Binary, female == 1, male == 0
Rural women	Independent	Gender, intersectional	
Self-employed women	Independent	Gender, intersectional	Binary, yes == 1, no == 0
Help needed	Dependent variable	Challenge	Binary, help needed == 1, no help needed == 0
Level of challenge experienced	Dependent variable	Challenge	Continuous, 1 to 5 (1 = very easy, 5 = very difficult)
App wouldn't open or has glitches	Dependent variable	Challenge	Binary, yes == 1, no == 0
No loan or credit offer	Dependent variable	Challenge	Binary, yes == 1, no == 0
Received offer, but never received monies	Dependent variable	Challenge	Binary, yes == 1, no == 0
No reminders for loan payments	Dependent variable	Challenge	Binary, yes == 1, no == 0
Poor customer service	Dependent variable	Challenge	Binary, yes == 1, no == 0
Any challenge	Dependent variable	Challenge	Composite indicator combining app wouldn't open or had glitches, no loan or credit offer, never received monies, no reminders for payments, poor customer service
Understood loan approval process	Dependent variable	Understandings (linked to challenge)	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Understood all terms and conditions	Dependent variable	Understandings (linked to challenge)	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
No trouble understanding how to make loan payments	Dependent variable	Understandings (linked to challenge)	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
No trouble affording	Dependent	Understandings	Continuous, 1 to 5 (1 =

payments	variable	(linked to challenge)	completely disagree, 3 = neutral, 5 = completely agree)
Understood how the Lender uses personal data	Dependent variable	Understandings (linked to challenge)	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Trust Lender with personal information	Dependent variable	Understandings (linked to challenge)	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Self-confidence	Dependent variable	Impact	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Hopefulness about the future	Dependent variable	Impact	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Sense of control over how money is spent	Dependent variable	Impact	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Ability to make financial decisions	Dependent variable	Impact	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Ability to meet basic needs	Dependent variable	Impact	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Financial well-being	Dependent variable	Impact	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)
Involvement in household decisions for large purchases	Dependent variable	Impact	Continuous, 1 to 5 (1 = completely disagree, 3 = neutral, 5 = completely agree)

Appendix 3, Table 2. Control and mediator variables

	Variab le type	**Category**	**Coding / Processing**	**Composite indicator (dimensions)**	
Age (youth)	Control	Demographi c	1 = under 35 0 = 35+	Combines 18-24 and 25-34	
Relationship status (married)	Control	Demographi c	1 = married 0 = not married	N/A	

Education (university)	Control	Demographic	1 = university degree 0 = no university education	N/A	
Income source (employed)	Control	Demographic	1 = employed 0 = not formally employed	N/A	
Income source (self-employed)	Control	Demographic	1 = self-employed 0 = not self-employed	N/A	
Location (rural)	Control	Demographic	1 = rural residence 0 = urban or peri-urban	N/A	
Personal phone used	Control / mediator	Digital / financial inclusion	1 = used personal phone to access the app 0 = used another person's phone	N/A	
Only Lender's app for digital loans	Control / mediator	Digital / financial inclusion	1 = only uses Lender's app for digital loans 0 = uses others in addition	N/A	
0 - 1 loans taken	Control / mediator	Digital / financial inclusion	1 = first-time borrower 0 = Taken 1+ loans	N/A	
Formal financial inclusion	Control / mediator	Digital / financial inclusion	1 = Formally financially included 0 = Formally financially excluded	Combines having received a loan from a formal financial institution, having opened a savings account with a formal financial institution, or having taken out formal insurance	
Other loan options	Control / mediator	Digital / financial inclusion	1 = Has other loan options 0 = No other loan options	N/A	

Loan purpose work	Mediator	Loan purpose	1 = Purpose of loan is for work / productive 0 = Purpose of loan is not for work / productive	Combines the latest loan being taken for starting a business or side hustle, buying inventory / stock / supplies, or farming supplies or resources	

Appendix 4. Correlation table

Appendix 4, Figure 1. Variance Inflation Factor (VIF) Table

This table presents the VIF values for all independent variables included in the regression models. All values fall well below the common threshold of 5, indicating no severe multicollinearity concerns.

Feature	VIF
const	11.78
Female	1.68
Married	1.42
Employed	1.43
Rural location	1.93
University degree	1.14
Youth (ages 18-24)	1.49
Older (Age 55+)	1.05
Exclusive use of lender app for digital loans	1.15
App loans 0 to 1	1.13
Financial inclusion	1.17
Loan purpose for work	1.17
Rural female	2.14
Self-employed female	1.58

Appendix 4, Figure 2. Correlation Matrix

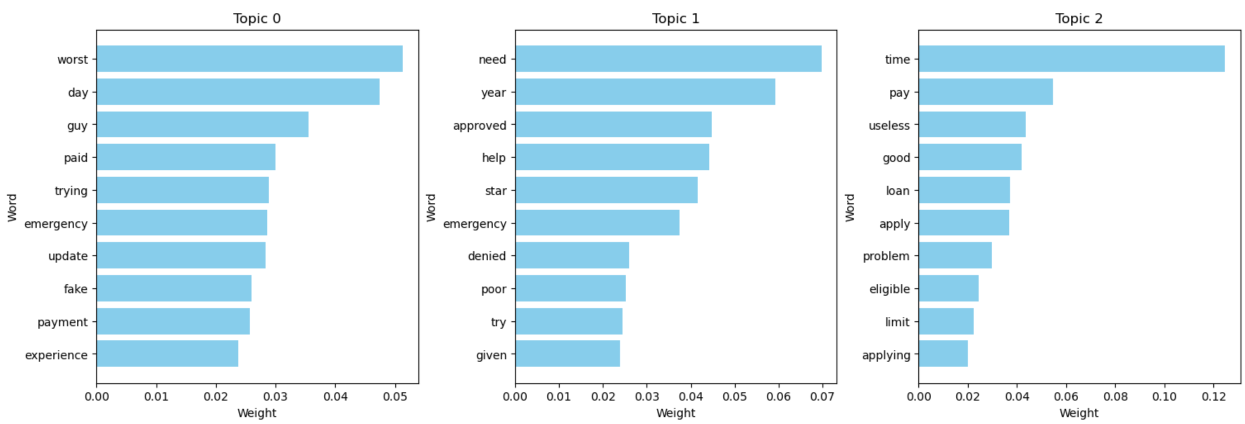
This correlation matrix shows pairwise Pearson correlation coefficients among the independent variables. While most correlations are modest, the highest correlation (0.66) is between rural location and the interaction term rural female, which is expected given its derivation. The matrix supports the VIF results in demonstrating low multicollinearity.

	Female	Married	Employed	Rural location	University degree	Ages 18-24	Ages 55+	Exclusive use of lender app	0 - 1 loans accessed	Financial inclusion	Rural female	Self-employed female	Loan purpose for work
Female	1.00	-0.29	-0.13	0.00	0.02	0.15	-0.01	0.19	-0.02	-0.12	0.28	0.46	-0.10
Married	-0.29	1.00	0.27	0.06	-0.01	-0.45	0.05	-0.09	-0.18	0.15	-0.11	-0.11	0.04
Employed	-0.13	0.27	1.00	-0.04	0.10	-0.25	-0.06	-0.25	-0.19	0.08	-0.10	-0.35	-0.17
Rural location	0.00	0.06	-0.04	1.00	-0.10	-0.10	-0.01	0.06	-0.03	-0.01	0.66	0.02	0.09
University degree	0.02	-0.01	0.10	-0.10	1.00	-0.05	-0.18	-0.10	0.02	0.25	-0.09	-0.01	0.01
Ages 18-24)	0.15	-0.45	-0.25	-0.10	-0.05	1.00	-0.08	0.10	0.23	-0.23	-0.09	-0.05	-0.17
Ages 55+	-0.01	0.05	-0.06	-0.01	-0.18	-0.08	1.00	0.08	0.00	-0.01	0.03	0.03	0.01
Exclusive use of lender app	0.19	-0.09	-0.25	0.06	-0.10	0.10	0.08	1.00	0.17	-0.12	0.10	0.08	0.02
0 - 1 loans accessed	-0.02	-0.18	-0.19	-0.03	0.02	0.23	0.00	0.17	1.00	-0.05	0.02	0.01	-0.06
Financial inclusion	-0.12	0.15	0.08	-0.01	0.25	-0.23	-0.01	-0.12	-0.05	1.00	-0.00	-0.04	0.18
Rural female	0.28	-0.11	-0.10	0.66	-0.09	-0.09	0.03	0.10	0.02	-0.00	1.00	0.15	0.07
Self-employed female	0.46	-0.11	-0.35	0.02	-0.01	-0.05	0.03	0.08	0.01	-0.04	0.15	1.00	0.18
Loan purpose for work	-0.10	0.04	-0.17	0.09	0.01	-0.17	0.01	0.02	-0.06	0.18	0.07	0.18	1.00

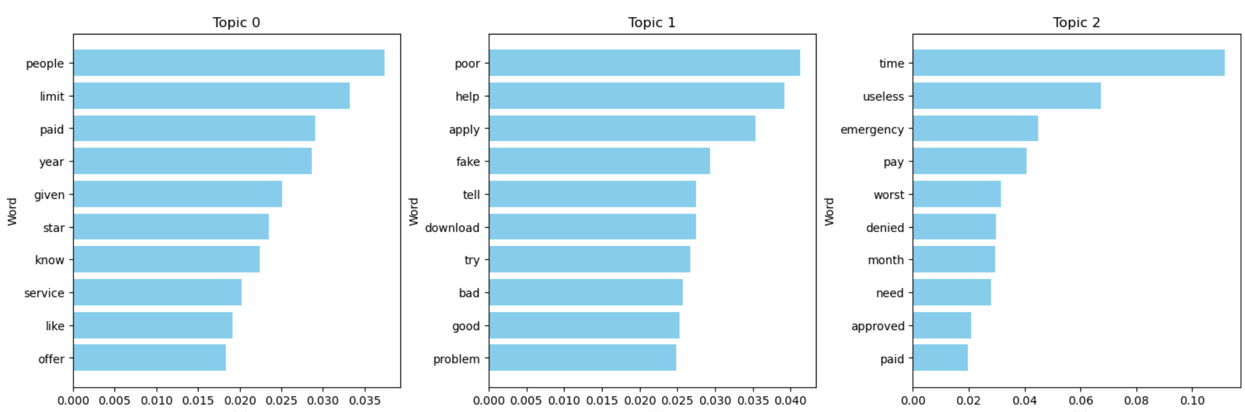
Appendix 5. Topic modeling

Topic modeling on Google Play review data among negative sentiment resulted in insights about issues experienced by borrowers. I identified three main areas of complaints amongst women: (0) payment issues and attempts, (1) loan denial and urgent financial needs, (2) application issues and time (see Appendix 4, Figure 1). The three primary topic areas for men include: (0) loan limits and customer service complaints; (1); application and download issues and frustration (2) denial and frustration with time and access (see Appendix 4, Figure 2).

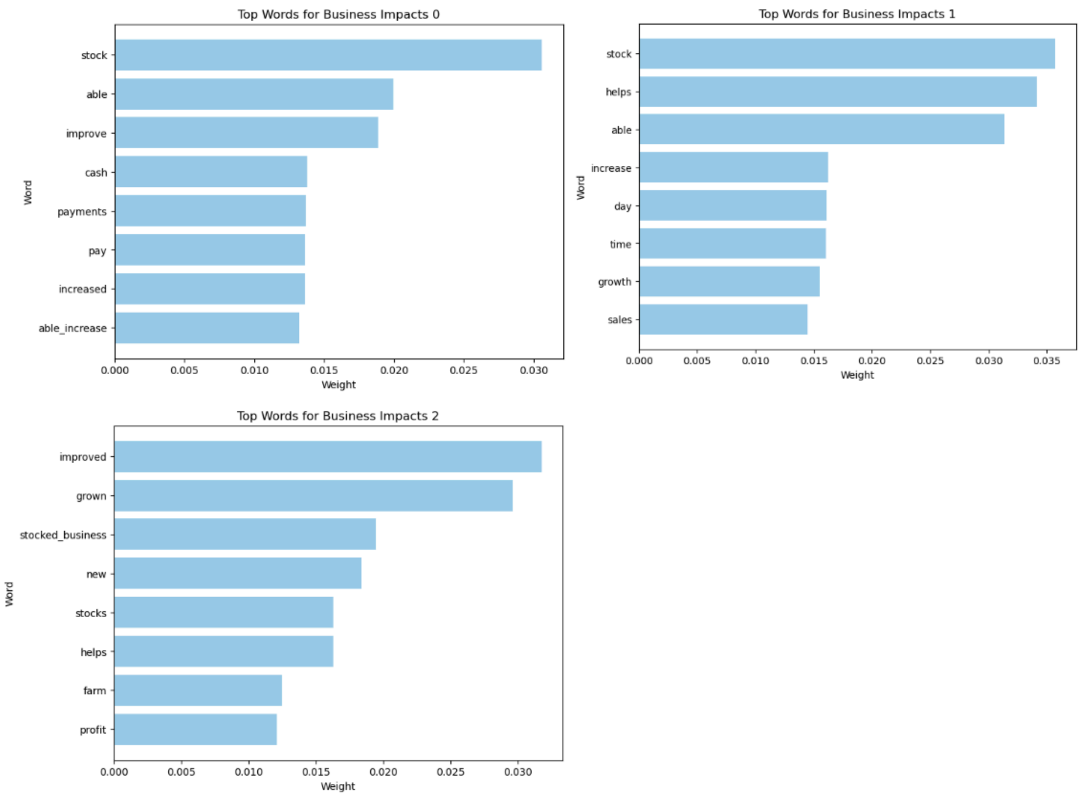
Appendix 5, Figure 1. Topic modeling of reviews with negative sentiment (<0), female



Appendix 5, Figure 2. Topic modeling of reviews with negative sentiment (<0), male



Appendix 5, Figure 3. Topic modeling results for open-ended response on impacts on business



Appendix 6. Understandings of the loan process

Appendix 6, Table 1. Understandings of the loan process, overall and by gender

Understanding	Overall (Mean [SD])	Female (Mean [SD])	Male (Mean [SD])	p (Gender)	Sig (Gender)	p (vs 3.0)	Sig (vs 3.0)
Understanding of loan approval process	4.45 [1.07]	4.49 [1.08]	4.41 [1.07]	0.5006	No	0.0000	***
Understanding of terms and conditions	4.48 [0.95]	4.55 [0.92]	4.42 [0.98]	0.2058	No	0.0000	***
No trouble understanding payments	4.61 [0.93]	4.57 [0.99]	4.65 [0.88]	0.4482	No	0.0000	***
No trouble affording payments	4.23 [1.18]	4.19 [1.25]	4.27 [1.11]	0.5192	No	0.0000	***
Understanding of personal data use	4.10 [1.18]	4.08 [1.22]	4.12 [1.14]	0.7539	No	0.0000	***

One-sample *t*-tests compare group means to a neutral value of 3.0 on a 1 to 5 scale for overall significance.
Chi-square tests assess whether reporting rates differ significantly by gender.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix 7. Interview Participant Information Sheet

DEPARTMENT CONTACT DETAIL
Genevieve Smith
genevieve.smith@kellogg.ox.ac.uk

DPhil student
Oxford University telephone number: +44 1865 281800
Oxford University e-mail: communications@qeh.ox.ac.uk

An examination of machine learning-based alternative lending on financial inclusion & gender equity
PARTICIPANT INFORMATION SHEET
Central University Research Ethics Committee Approval Reference: C1A_23_079

1. **Introductory paragraph**

You are being invited to take part in a research project. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Take time to decide whether you wish to take part and ask any questions for clarification or further information.

2. **Why is this research being conducted?**

The aim of this research is to investigate whether and how machine learning-based alternative lending tools by fintechs impact financial inclusion and gender equity in low and middle income countries. We appreciate your interest in participating in this discussion.

3. **Why have I been invited to take part?**

You have been identified as a potential participant in the study given your affiliation with a fintech company that is developing and utilising machine learning-based alternative lending. You were chosen based on your leadership in this space and position at the company, with 23 other participants being recruited for the study.

Inclusion criteria for the research includes: (1) being an employee (either management leader and/or data scientist leader) at a fintech company that develops and/or manages an machine learning-based alternative lending tool assessing creditworthiness using alternative data in low-and middle-income countries and has been in operation/existence for at least the past 12 months; (2) speaks English; (3) is over 18 years old. Exclusion criteria includes (1) the fintech of the employee being in operation/existence for less than 12 months; (2) the participant does not speak English; and (3) the participant is under 18 years old.

4. **Do I have to take part?**

No. It is up to you to decide whether to take part. You can withdraw yourself from the research, without giving a reason, by advising me/ us of this decision. The deadline by which you can withdraw any information you have contributed to the research is before data is anonymized. If you decide to withdraw, any data already collected will be destroyed and not used.

5. **What will happen to me if I take part in the research?**

- The one interview will take place virtually via Microsoft Teams.

- Consent will be taken through a written consent form prior to the interview.
- The interview will be 45-60 minutes.
- The interview will include a range of questions about the background of the fintech, the conceptualization of the machine learning tool, design of the tool, and management of the tool. It will also include questions related to the demographics of those involved in building and managing the tool.
- With your consent, I would like to audio record you so I can have an accurate record of our conversation.
- You can ask to pause or stop the interview at any time.

6. What are the possible disadvantages and risks in taking part?

There are no foreseeable discomforts, disadvantages or risks to you. All data from interviews will be anonymous and confidential through removal of names and unnecessary personal details from interview transcripts. Any proper names will be changed to codes. In the case of utilising quotes directly attributable to you, the quote will be shared with you for review with a request for explicit written consent. Then, pseudo-anonymized data will be used (e.g., a data scientist of one fintech in the United States).

7. Are there any benefits in taking part?

While there are no immediate benefits for those people participating in the research, it is hoped that this research will lead to a better understanding of the ways that machine learning-based credit assessment tools impact financial inclusion and gender equity, as well as opportunities to enhance financial inclusion and gender equity.

8. What information will be collected and why is the collection of this information relevant for achieving the research objectives?

I am interested in perspectives and approaches towards using machine learning in alternative lending, and impacts and learnings related to financial inclusion and gender equity. I am interested in your experience as a leader at a fintech organisation and in the field more broadly. The information you provide will help me better understand the ways that such tools are conceptualised, designed, and managed to answer my research question on whether and how machine learning-based alternative lending tools by fintechs impact financial inclusion and gender equity in low and middle income countries. Data from the interview will be password protected and stored in OneDrive for Business, which is approved by the University's Information Security team for the secure storage of research data. The researcher and supervisors (Prof. Xiaolan Fu and Prof. Masooda Bano) will have access to the research data. Identifiable data (including consent forms) will be password protected and stored in OneDrive for Business. Three years after the end of the research project, all data will be destroyed.

9. Will the research be published? Could I be identified from any publications or other research outputs?

The findings from the research will be written up in a thesis dissertation. Findings may also be written up for academic publications and conference presentations. Participants will not be identifiable from the outputs. A copy of my thesis/ dissertation will be deposited both in print and online in the Oxford University Research Archive where it will be publicly available to facilitate its use in future research.

10. Data Protection

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that is performed in the public interest. Further information about your rights with respect to your personal data is available from the University's Information Compliance web site at https://compliance.admin.ox.ac.uk/individual-rights.

11. Who has reviewed this research?

This research has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee. (Ethics reference: C1A_23_079**)**.

12. Who do I contact if I have a concern about the research or I wish to complain?

If you have a concern about any aspect of this research, please contact Genevieve Smith (genevieve.smith@kellogg.ox.ac.uk) or Xiaolan Fu (xiaolan.fu@qeh.ox.ac.uk), and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:
Prof. Loren Landau, The Chair, Department of International Development Research Ethics Committee; Email: loren.landau@qeh.ox.ac.uk; Address: 3 Mansfield Rd, Oxford OX1 3TB, UK

13. Further Information and Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Genevieve Smith
Department of International Development
3 Mansfield Rd, Oxford OX1 3TB, United Kingdom
University tel: +44 1865 281800
University email: communications@qeh.ox.ac.uk

Appendix 8. Interview Participant Consent Form

DEPARTMENT CONTACT DETAIL
Genevieve Smith
genevieve.smith@kellogg.ox.ac.uk

DPhil student
Oxford University telephone number: +44 1865 281800
Oxford University e-mail: communications@qeh.ox.ac.uk

OXFORD DEPARTMENT OF INTERNATIONAL DEVELOPMENT

UNIVERSITY OF OXFORD

Consent to take part in *An examination of machine learning-based alternative lending on gender equity*

Central University Research Ethics Committee (CUREC) approval reference: C1A_23_079

Purpose of Study: The purpose of the study is to investigate whether and how machine learning-based alternative lending tools impact financial inclusion and gender equity in low and middle income countries.

	Please initial each box if you agree with the statement
I confirm that I have read and understand the information sheet for the above research. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
I understand that my participation is voluntary and that I am free to withdraw at any point until data is anonymised, without giving any reason.	
I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	
I understand that I will not be identifiable from any publications or in the thesis.	
I consent to being audio recorded.	

I understand how audio recordings will be used in research outputs. ☐

I understand In the case of utilising quotes directly attributable to me, the quote will be shared with me for review with a request for explicit written consent. Then, pseudo-anonymized data will be used (e.g., a data scientist in the United States). ☐

I give permission for you to contact me again to clarify information. ☐

I understand how to raise a concern or make a complaint. ☐

I agree to take part. ☐

Name of participant: ______________________ Date: dd / mm / yyyy

Signature: ______________________

Name of person taking consent: ______________________ Date: dd / mm / yyyy

Signature: ______________________