MIS Quarterly

RESEARCH ARTICLE

PROFIT VS. EQUALITY? THE CASE OF FINANCIAL RISK ASSESSMENT AND A NEW PERSPECTIVE ON ALTERNATIVE DATA[1]

Tian Lu
Department of Information Systems, W. P. Carey School of Business,
Arizona State University, Tempe, AZ, U.S.A. {lutian@asu.edu}

Yingjie Zhang
Department of Marketing, Guanghua School of Management,
Peking University, Beijing, CHINA {yingjiezhang@gsm.pku.edu.cn}

Beibei Li
Heinz College of Information Systems and Public Policy,
Carnegie Mellon University, Pittsburgh, PA, U.S.A. {beibeili@andrew.cmu.edu}

The importance of pursuing financial inclusion to accelerate economic growth and enhance financial sustainability has been well noted. However, studies have provided few actionable insights into how financial institutions can balance the potential socioeconomic trade-off between profitability and equality. One major challenge arises from a lack of understanding of the impacts of various types of market information available on financial equality beyond economic profitability. Another challenge lies in how the socioeconomic trade-off under a large set of counterfactual policies in a real-world setting can be evaluated. Our motivation for the present study was the emerging sources of digitized user-behavior data (i.e., "alternative data") stemming from the high penetration of mobile devices and internet access. Accordingly, we investigated how alternative data from smartphones and social media can help mitigate potential financial inequality while preserving business profitability in the context of financial credit risk assessment. We partnered with a leading microloan website to design a novel "meta" experiment that allowed us to simulate various real-world field experiments under an exhaustive set of counterfactual policies. Interestingly, we found that profiling user financial risk using smartphone activities is 1.3 times more effective in improving financial inclusion than using online social media information (23.05% better vs. 18.11%), and nearly 1.3 times more effective in improving business profitability (42% better vs. 33%). Surprisingly, we found that using consumers' online shopping activities for credit risk profiling can hurt financial inclusion. Furthermore, we investigated potential explanations for financial inclusion improvements. Our findings suggest that alternative data, especially users' smartphone activities, not only demonstrate higher ubiquity but also appear to be more orthogonal to conventional sensitive demographic attributes. This, in turn, can help mitigate statistical bias driven by the unobserved factors or underrepresentative training samples in machine-based risk assessment processes.

Keywords: Credit risk, alternative data, financial trade-off, financial inclusion, profitability, equality, biases, counterfactual simulated experiment

[1] Gal Oestreicher-Singer was the accepting senior editor for this paper. Lynn Wu served as the associate editor.

Introduction

Financial technology (*fintech* hereafter), with its increasing popularity, has expanded access to finance for millions of people and has enabled society to promote inclusive finance (Guild, 2017). It has achieved these things specifically by offering microcredit to low-income individuals and budding entrepreneurs who have difficulty obtaining financial support from traditional financial institutions such as banks (Lu et al., 2020).

Since reaching out to a broader scope of users has great potential to expand the consumer pool (i.e., to improve commercial sustainability), the concept of financial inclusion is intriguing for all kinds of financial institutions (Davis, 2021). Indeed, financial inclusion has always been considered to have positive outcomes for society in terms of economic growth and financial equality (Sethi & Acharya, 2018; Berentsen & Markheim, 2021; Maskara et al., 2021). It has been reported that by opening the doors of financial services to many more people, fintech could increase the GDP of emerging economies by as much as 6% (or $3.7 trillion) by 2025.[2] Since up to 1.7 billion people globally lack access to formal financial services, there is wide latitude for efforts to boost inclusive finance (Loufield et al., 2018).

However, whereas the importance of promoting inclusive finance to accelerate economic growth and financial sustainability has been well noted (Loufield et al., 2018), prior work has identified the significant challenges to balancing the potential socioeconomic trade-off between financial inclusion and business profitability (e.g., Cull et al., 2011; Bassem, 2012). For example, Cull et al. (2011) revealed that the cost to financial institutions would be much higher if they made smaller loans on average and served more women. While offering more loans or targeting applicants from disadvantaged backgrounds is a simple strategy to increase financial equality, it incurs higher risk and may reduce profitability. Therefore, microfinance institutions seeking only to maximize revenue generally favor less-risky users with higher incomes or more education—undermining the goals of financial equality (Óskarsdóttir et al., 2019). Unfortunately, to date, very few actionable recommendations on how financial institutions should balance the potential socioeconomic trade-off between profitability and equality have been proposed—mostly due to a lack of understanding of the impact of various types of information on financial equality *beyond* economic profitability. Moreover, any actionable recommendations would require *real-world causal* evaluation under an alternative (and large) set of counterfactual policies.

[2] https://www.euromoney.com/article/b12kq8l0pmkrw7/financial-inclusion-fintech-and-the-gdp-boost.

Our motivation for this study was the emerging sources of digitized user-behavior data (i.e., "alternative data") stemming from the high penetration of mobile devices and internet access. In particular, the present study aimed to answer the following research questions: (1) *What kinds of alternative data can help balance the trade-off between financial profitability and equality?* (2) *What are the potential explanations?*

We investigated how alternative data from smartphones and social media can help mitigate potential financial inequality while preserving business profitability in financial credit risk assessment. In contrast to conventional data that typically originate from a credit bureau, a credit application, or a lender's own records on existing consumers, alternative data is most frequently obtained from public social media sites or individual applications and devices. Alternative data is a new and unparalleled source of fine-grained user behavior information that can be obtained at a reasonable acquisition cost (Tan et al., 2016). Although such data might not directly relate to a consumer's financial credit behavior, it has significant potential to complement conventional data in enhancing credit risk assessment accuracy and thus profits (Carroll & Rehmani, 2017).

More importantly, we believe that the great value of alternative data, in the present context, lies in its potential to improve financial inclusivity while minimizing financial losses. Rich sources of alternative data offer timely remedies for financial institutions seeking to cover the "thin-file" population (Loufield et al., 2018), and leveraging such new sources of information may reduce unintended bias in the risk evaluation process. A major source of unintended bias is the unobserved attributes of loan applicants (Dobbie et al., 2018). That is, traditional credit risk assessment relies heavily on a few (sensitive) demographic features, which might not capture individuals' real-time financial status and psychological assets (e.g., willingness to repay loans on time). When extracted from rich sources, alternative data can enable a broader understanding of individual attributes, including those that might go unnoticed in conventional models.

Therefore, if novel features from alternative data that are predictive of user financial risk but orthogonal to sensitive demographic attributes can be identified, financial companies could theoretically increase their profits and simultaneously mitigate financial inequality. *This represents one goal of the present study*. Some types of individual behavior might present similar distributions across different demographic subgroups, while others might appear to be more discriminative. For example, cellphone call frequency might be especially useful for evaluating credit risk, as it reflects a user's social network status (Ma et al., 2018) and is indifferent between men and women (Roberts et al., 2014). Hence, features such as cellphone call frequency can be exceedingly valuable because they are not only predictive of individual financial risk but also are relatively independent of sensitive attributes. By contrast, features such as excessive alcohol consumption, which reflects a lack of self-discipline that potentially predicts creditworthiness, are observed more frequently in men than in women (Luchetti et al., 2018). Thus, although observations of alcohol consumption can improve risk assessment and boost profits, this feature may also deepen the financial service accessibility rift between men and women.

Meanwhile, another unintended source of inequality is training sample bias (e.g., Cowgill et al., 2020). Financial companies use training data heavily biased toward successfully approved loan applicants ("approved samples" hereafter) whose credit risk had been perceived to be low enough for loan approval.[3] These approved samples tend to have lower default probabilities and significantly different socioeconomic characteristics (e.g., higher income, better educated) compared to the true population of loan applicants. The patterns or relationships learned from such biased samples might have limited generalizability and may hence lead to poor predictive performance for new applicants. This would especially be the case for applicants whose characteristics (e.g., lower income, less educated) may have barely registered in previous training samples. Worse, if initially approved samples are biased (intentionally or unintentionally) toward certain sensitive attributes, such errors could be further amplified when training with approved samples (Fu et al., 2021). Therefore, *another goal of this study* was to examine training sample bias in the context of financial risk assessment in order to understand how alternative data can help mitigate such bias.

To date, both industry and academia have been seeking other and various sources of data to improve credit risk prediction (e.g., Ma et al., 2018; Jagtiani & Lemieux, 2019). However, the literature lacks answers regarding whether and, if so, what kinds of alternative data can benefit financial profitability and financial equality simultaneously and why. To achieve our goals, we partnered with a microloan website in Asia to conduct a large experiment. Instead of randomly assigning applicants to various subgroups with different selection algorithms, we designed a novel "meta" experiment wherein all applicants were approved without any selection criteria during the experimental period. We then gathered all behavioral data (including individual characteristics, loan features, and corresponding repayment histories) from this entire applicant population. By approving all loan applications and tracking borrowers' repayment behaviors over time, we were able to observe counterfactual cases—applicants who, in normal circumstances, would have been rejected. Moreover, we were able to observe *all* possible counterfactual cases by simulating all possible treatments (i.e., different loan-approval ranking mechanisms) and tracking the corresponding loan repayment behaviors and financial outcomes. In essence, our "meta" experiment allowed us to simulate various real-world field experiments under an exhaustive set of counterfactual controlled conditions.

The key strength of this novel "meta" experimental design—especially when compared to the traditional A/B type of field experiment—lies in the fact that the loan approval decision is clean, exogenous, and algorithm independent because all applicants are granted loans. Thus, for the first time in this research context, there was no sample selection, and the counterfactuals for any simulated algorithm using any data source could be observed. By including behavioral patterns from the entire loan applicant population, this unique setup also enabled us to accomplish two things: form an unbiased training sample for model training and evaluate our model under various counterfactual scenarios that otherwise would have gone unobserved.

Our data cover multiple information sources. Inspired by the existing literature, we constructed and extracted more than 100 features spanning four categories: commonly adopted conventional data, online activities (e.g., shopping), mobile activities (e.g., cell phone usage and location mobility traces), and social media activities. We first subjected those features to a machine learning training process to determine individual applicants' credit risk. We then disentangled the value of alternative features in overcoming the trade-off between financial profitability and equality by empirically calculating both the economic gains to the microloan website and the boost in financial equality.

Our empirical analysis yielded several interesting findings. First, a welfare analysis with a fixed interest rate scheme (i.e., the company randomly offered yearly interest rates within a narrow range to borrowers, unrelated to their evaluated credit risk) indicated that all of the proposed alternative feature sets can boost profits. In particular, when predicting borrowers' credit risk according to their mobile activities, the corresponding loan selection strategy yielded 22% more revenue for the microloan website than using conventional

[3] Applications initially perceived to be high risk, on the other hand, tend to be immediately rejected, with the result that no further loan payment data on these applicants are recorded or included in further model training.

features only. Our analyses also suggested guidelines for the design of proper selection strategies under different company budget constraints. Second, in terms of financial equality, we found that the existing approach using conventional features only tends to favor higher-income and more-educated applicants from more economically developed areas. We further discovered that by leveraging alternative data from smartphone usage and social media, the focal microloan website was more likely to include lower-income and less-educated loan applicants from less-developed geographic areas—which is to say, historically disadvantaged and largely neglected populations. Our study thus demonstrates the tremendous potential of leveraging these alternative data to alleviate financial inequality while simultaneously achieving higher revenues. Third, and more interestingly still, a finer-grained analysis by which we were able to decompose the value of different types of alternative features revealed that alternative data from mobile activities are more effective than data from online social media in balancing the trade-off between financial profitability and equality. We observed that profiling users' financial risk using smartphone activities is almost 1.3 times more effective than using online social media information in improving financial inclusion (23.05% better vs. 18.11%) and nearly 1.3 times more effective in improving business profitability (42% better vs. 33%). Surprisingly, observations of online shopping activities did not correlate with improved financial inclusivity. A mechanism analysis showed that this was due mostly to the high correlation between such activities and certain sensitive user attributes.

Next, we took a further step in empirically testing the training sample bias. We first demonstrated that the bias indeed existed when using either only approved samples or only conventional data. In both cases, significant losses of prediction accuracy and economic profit resulted. Nevertheless, we observed that even with sample bias, the economic gains of applying multiple sources of data (15,410 USD) were much larger than when simply applying conventional data without sample bias (13,920 USD). These findings indicate that alternative data can help shrink the economic losses caused by sample bias and that the economic value exceeds that obtained using conventional data only, even without sample bias. More importantly, our feature-importance analyses revealed that alternative features can better capture borrowers' credit risk from different angles that are orthogonal to sensitive demographic features such as gender and income. Therefore, a predictive model based on alternative features is less likely to cause a bias toward sensitive attributes.

The contributions of our study are multifold. First, this study extends the microlending literature by focusing on a solution to the problem of balancing the trade-off between financial profitability and equality. It validates an actionable scheme that takes advantage of alternative data to balance that trade-off. Specifically, this study is among the first to identify different types of alternative data that contribute to improvements in financial inclusion in the microloan industry. Second, this is the first study to systematically investigate the economic value of multidimensional alternative data (including cellphone and mobile app usage, mobility trajectories, shopping behavior, and social media information). We also separately identify individuals' delinquent and delinquent-but-not-in-default behavior in the microloan setting. This extra step allowed us to offer insights into how proper loan selection strategies could be designed under company budget constraints. Third, for evaluation, we used a unique experimental setting to examine "what-if" counterfactual scenarios under different loan selection strategies. Fourth, we offer an approach whereby microloan companies can easily adopt cost-effective solutions based on what is easier to implement in practice. For example, training sample bias represents a major challenge in both prior research and industry practice due to practical data limitations. We demonstrate that incorporating alternative data can not only largely offset the potential economic loss caused by training sample bias; it can also lead to significant improvement in revenues, even when microloan companies have no access to the unbiased full sample of loan applicants.

Literature Review

The Trade-Off between Financial Profitability and Equality

A few studies have paid attention to the potential trade-off between financial profitability and equality. For example, Bassem (2012) showed the contradiction between good financial profitability and a high depth of outreach (serving the poor) using a sample of 64 microfinance institutions from the Middle East and North Africa regions. Churchill (2020) confirmed this trade-off between financial performance and outreach depth using data on 1,595 microfinance institutions (MFIs) in 109 countries. Additionally, Cull et al. (2011), compared nonprofit MFIs with commercialized microfinance banks, pointing out that the former had to bear more cost per loan dollar if they made far smaller loans on average and served more women as a fraction of their total customers. Based on industry-level and cross-country data, the above literature, without exception, has demonstrated that the trade-off between financial profitability and equality does in fact exist. In contrast, very few studies have focused on solutions, especially those from the perspective of a single company. In light of this, the present study aims to investigate and propose a useful and actionable device to help financial instructions balance the trade-off in their business operations.

Microlending, Fintech, and Financial Inclusion

Most of the current microlending literature has focused on loan selection strategies primarily aimed at screening applicants' credit risk, especially their default rate. One stream of research has focused on comprehensive data and features relevant to microloan or peer-to-peer (P2P) lending. Scholars and practitioners have investigated the usefulness of conventional features such as loan characteristics, borrower characteristics, credit history, and social capital (e.g., Iyer et al., 2016; Lin et al., 2013; Liu et al., 2015; Serrano-Cinca et al., 2015; Wei et al., 2015). With greater access to alternative data such as cellphone usage (e.g., Tan et al., 2016; Mehrotra et al., 2017; Ma et al., 2018), social media information (e.g., Ge et al., 2017; Tang, 2019), and employment data (Chan et al., 2020) in recent years, several scholars have revealed the value of these data for default risk prediction. In addition to the above accuracy-oriented prediction analyses, a few researchers have recently turned their attention to profit-based analyses (e.g., Serrano-Cinca & Gutiérrez-Nieto, 2016; Papouskova & Hajek, 2019). These studies mostly predicted the expected profitability of investing in P2P loans and/or optimized profit-based models.

Another inherent goal of microfinance is to realize financial inclusion, or "the delivery of financial services at an affordable cost to vast sections of disadvantaged and low-income groups" (Dev, 2006). Providing access to financial services has significant benefits. It allows people to make financial transactions more efficiently and safely and to better manage financial crises (Demirguc-Kunt et al., 2017). More importantly, it diminishes financial inequality and helps lift the poor out of the cycle of poverty by providing growth opportunities that facilitate their consumption and investment in health, education, and income-generating activities (Yawe & Prabhu, 2015).

With the introduction of fintech, microlending companies can now use nontraditional alternative sources to collect soft information about the creditworthiness of individual users and small business owners—especially those with little or no credit history—and to apply the collected information to big data analytics (Jagtiani & Lemieux, 2019). This helps to expand microlending services to lower-income users and "thin-file" users who may otherwise be excluded from the formal or traditional financial sector (Loufield et al., 2018). Some recent studies have observed that the use of alternative data has enabled some borrowers who would have been classified as subprime by traditional criteria to be slotted into "better" loan grades, which allowed them to get lower-priced credit (Jagtiani & Lemieux, 2019). The current study performs a thorough analysis of the types of data that can boost financial inclusion in microlending and how they are able to do this.

Biases in Decision-Making with Machine Learning

Irrespective of the pursuit (or not) of financial inclusion, prior studies have suggested that decision-making under uncertainty using machine learning may result in discriminatory or biased outcomes (Fu et al., 2022). Such biases stem from various sources—for example, the use of a limited number of observable sensitive features (e.g., gender and ethnicity) that correlate closely with both creditworthiness and observed input features (Chouldechova et al., 2018, Dobbie et al., 2018, Fuster et al., 2022). To deal with biases and mitigate unintended unfairness, recent studies have amended extant models and developed debiasing algorithms by using class attributes and/or accommodating group-specific thresholds (e.g., Berk et al., 2017; Fu et al., 2021). However, these strategic manipulations might not fully consider decision makers' intrinsic goals and generally suffer from drawbacks such as insufficient bias removal, lack of generalizability, or low accuracy (Cowgill & Tucker, 2019).

Further, in many cases, bias may also appear within the training data according to the way that outcomes are evaluated for users/applicants who have been approved (Cowgill, 2019). Decision makers' training data are often missing for some nonrandomly selected applicants. This is problematic when the goal is to develop an algorithm for use in screening a larger or full population (Cowgill & Tucker, 2019). In such cases, the distributions of the features of the selected applicants and the full population would be discrepant. Explicitly, the way to cope with this bias is to obtain a full sample for training or, at the very least, to use a carefully designed complementary model with additional experimentation, as suggested by Cowgill (2018).

One neglected area in credit risk prediction is samples. The previous studies are based mostly on approved "good" loans, which would incur bias when using training results to screen full applicants. Also, whereas previous studies have demonstrated the effectiveness of alternative data, no systematic comparison of predictive power among different sources of alternative data has been performed. Meanwhile, although Serrano-Cinca and Gutiérrez-Nieto (2016) and Papouskova and Hajek (2019) have focused on profit scoring, their proposed models have limited predictive power due to inadequate predictive features. This paper addresses all of these relevant problems.

Context, Experiment, and Data

Experimental Design and Setup

We collaborated with a medium-sized Asian microloan website that was founded in 2011. The website offers microloans at an average size of approximately 450 USD and has an accumulated transaction volume of more than 2 million USD. Loan periods range from one to seven months. The website uses only its own money for lending; its borrowers, meanwhile, use the loans primarily to support temporary financial needs, including supplementary working capital for small businesses, irregular shopping needs, education spending, and medical expenses. The website generates revenue from the interest paid by non-defaulting borrowers and the penalties paid by those paying in arrears (refer to Appendix B for details). The website's costs are incurred mostly from the unpaid principal of defaulting borrowers. Besides, the website collects a fixed commission fee from each borrower to balance the labor costs of credit risk evaluation and debt collection. The final default records are submitted to a centralized shared blacklist system maintained by a symposium of microloan companies. In cases of default, the website may take legal action.

The realization of both profit and financial inclusion for a company is predominantly reliant on its selection strategy, which is contingent on a combination of various feature sets, training models, and approval rates. Given a specific selection strategy (i.e., the treatment), only some loan applicants would be approved. Following the causal framework of the Roy-Rubin model (Roy, 1951; Rubin, 1974), let us denote $Y_i(1)$ as the potential outcome for individual i if the individual is treated (i.e., receives the loan in the given selection strategy), and $Y_i(0)$ as the potential outcome if the individual is not treated. Thus, given a treatment $D = 1$, the average treatment effect (ATT) in this framework could be denoted as $\pi_{ATT} = E[Y(1)|D = 1] - E[Y(0)|D = 1]$. Our treatment D is a specific loan selection strategy determined by the applied features, the machine learning algorithm, training samples, and behavioral indicators for prediction. A commonly used approach is using a randomized field experiment (or A/B test) to approximate $E[Y(0)|D = 1]$ with $E[Y(0)|D = 0]$. Such a design would allow us to measure the outcome of a single selection strategy. Therefore, determination of the optimal selection strategy ideally requires a typical experiment with multiple pairs of treated (i.e., applicants approved by the strategy) and control (i.e., applicants denied by the strategy) groups. However, this is difficult and highly costly to implement since we needed to compare a large number of strategies (with diverse feature sets, prediction models, samples, and behavioral indicators). Therefore, we needed an alternative strategy, other than a typical randomized field experiment, to achieve our goal. Given the unique setup of our context, wherein the control group is comprised of applicants whose loan applications were not approved, such applicants would produce zero profit for the company (because $Y_i(0) \equiv 0$ for rejected loan applications).

This inspired us to address the challenge by conducting an algorithm-independent experiment wherein all applicants are approved without any selection strategy. In this case, we observed $Y_i(1)$ for any individual i. In sum, with this "meta" experimental design, we observed both $Y_i(0)$ and $Y_i(1)$ for every loan applicant, which allowed us to perform counterfactual comparisons and simulate the performance of different strategies. Given that our meta-experiment allowed us to observe individual-level treatment effects using $Y_i(1)$ and $Y_i(0)$ for any individual i, we were able to construct different causal estimates by integrating the pool of individuals selected by a certain treatment D. Further, because our experimental design allowed us to observe potential outcomes for each individual, we were able to determine individual treatment effects (Heckman & Vytlacil, 2007).

We conducted our experiment on the focal microloan website from December 2 to December 22, 2017. During the experimental period, the website randomly selected 40% of all loan applicants and approved all of them, without using any selection strategies. This experimental design allowed us to keep track of the repayment behaviors of all of the borrowers (full samples). Since our experiment was based on fixed interest rates, the company randomly offered yearly interest rates within a narrow range (12% - 16%) to borrowers (see Table A2 for details).[4] While this fixed interest rate scheme would affect the estimates of profitability to some extent, it would not influence the ability to predict credit risk under diverse scenarios in our analyses. Thus, it allowed us to identify the value of alternative data without interference from the interactions of interest rate and alternative data (via predictive ability).

Data

The full experimental sample, with three parts, had 5,214 loans granted to 5,214 unique borrowers.

Conventional information: For each loan, we collected the following information: (1) loan attributes (i.e., loan amount, loan term, and interest rate); (2) borrower's demographic and socioeconomic characteristics (i.e., age, gender,

[4] The interest rate was disclosed to the borrower in the final step of the loan application and borrowers had no prior information on the specific interest rate they would finally get. Namely, the self-selection issue was less of a concern. Since the loan amount and interest rates overall are small, the manager confirmed that only six borrowers abandoned their loan grants after the disclosure of interest rates.

education level, income level, marital status, number of children, job, and contact information of at least one family member or close relation)[5]; (3) the borrower's self-reported purpose for the loan;[6] (4) the borrower's loan history on the focal and other microloan websites. These sources of information are commonly employed in credit risk assessment (Serrano-Cinca et al., 2015).

Alternative data: For each loan application, the website collects alternative sources of information covering the applicant's behavior during the six months prior to the application.[7] The website considers the following three types of alternative data: (1) online (shopping) activities records (i.e., order time, product name, price, quantity, product type, and receiver information) from the two largest online shopping platforms in the country; (2) mobile activities records (i.e., call history, cellphone usage, detailed mobile app usage, GPS mobility trajectories[8]); (3) social media activities records (i.e., whether the borrower has accounts and, if so, all posted messages with time stamps, social media presence including number of fans, followings, received comments, and received "likes") at the largest Twitter-like microblog community in the country (see Appendix A for an illustration of a microblog).

Repayment information: We collected the repayment behavior (i.e., due date and repayment rate) of each loan at the installment level (monthly). On the website, borrowers must repay installments every month until the loan is paid off. If a loan is not paid off three months after the due date, loan default is confirmed.

Analysis Setup

This section introduces our analysis setup based on the above data sources. Specifically, we first elaborate our feature extraction details, followed by our credit risk assessment (including both the evaluation outcomes and model training process).

Feature Extraction

We constructed and extracted 117 features covering the following four main categories: commonly adopted conventional characteristics, online (shopping) activities, mobile activities, and social media activities. We summarize the definitions and statistics in Appendix A.

Commonly adopted conventional characteristics (*Fc*): Concretely, for borrowers' demographic and socioeconomic characteristics, we coded for the following factors: age, gender, education level, marital status, number of children, homeownership, type of occupation, monthly income, and whether they have insurance. Since individual income level was reported by the borrowers themselves, we coded the disposable personal income (DPI) in 2017 of residents in the city the borrower lives in as a supplementary feature of income. As loan attributes, we used loan amount, loan period, interest rate, and income-to-debt ratio. We also coded three features indicating whether the installment payment due date was during a holiday, weekend, or the beginning/end of a month, respectively. For loan histories, we coded whether borrowers had microloan experience with the focal website or other microloan companies, whether they had defaulted on prior microloans, their frequency of contact with the microloan company, and whether they had credit cards and exhibited regular payment behavior. Additionally, we extracted borrowers' self-reported loan purposes via text mining techniques and coded them as a binary feature, indicating whether the loan was used for (high) consumption or emergencies (e.g., healthcare, accidents, or business turnover) (1 = *consumption*, 0 = *otherwise*). We also included a feature derived from the default behavior of borrowers' (first-order) recent cellphone contacts.

Online (shopping) activities (*Fo*): To some extent, online shopping behavior reflects borrowers' ability to pay as well as personal preferences. We filled in the overall picture of a shopper using multiple dimensions. We started with several general characteristics, including the total amount transferred out and in through a third-party payment website that has been widely accepted by e-commerce platforms. We then decomposed the borrower's shopping behavior by considering different types of products. For each type, we aggregated the total number of transactions, average amount and quantity of purchased products, and diversity of purchased product categories. We considered four types, including durable goods, game products (suggesting a self-indulgent intention— Kim et al., 2008), special products such as medicine, caffeine, and tobacco (Amonini & Donovan, 2005), and products purchased for others.

[5] Borrowers must provide the numbers of contact people and their relationships. The website validates the information by checking, based on call logs, whether or not the borrowers frequently interact with their contact people.

[6] As an additional analysis, we report credit risk predictions using different sets of variables across the two types of loan purposes in Appendix C.

[7] A successful loan application requires the borrower to authorize the focal microloan website to collect these alternative sources of personal information for credit risk screening.

[8] GPS-based location data points are recorded every two hours. Based on this geographical information, we extracted the appearance frequencies in different locations or trajectories of borrowers.

Mobile activities (*Fm*): Several studies have established a relationship between cellphone usage/mobility data and socioeconomic status (e.g., Blumenstock et al., 2015; Soto et al., 2011), friendship and social ties (e.g., Onnela et al., 2007), and psychological and personality traits (e.g., Kim et al., 2015). These sets of features cover three aspects. The first one covers call- and message-related features, including average monthly expenses, the frequency and duration of incoming and outgoing calls, the number of sent and received text messages, and the number of calls with the close relations the borrower registered during the loan application stage (Ruiz et al., 2017). Second, inspired by Ma et al. (2018), we extracted app usage and data traffic usage features. The app usage features include average weekly frequency and duration of different kinds of app usage (e.g., financial and payment apps, news apps, game apps, entertainment apps, and social media apps), and data traffic usage. The third subset captures an individual's offline trajectories (Tan et al., 2016) extracted from fine-grained GPS data. Specifically, we calculated the number of cities the borrower visited and their average weekly frequency of visiting different location types, including office buildings and commercial, entertainment/recreational, and public service places.

Social media activities (*Fs*): Our sample included 1,618 (31.03%) borrowers who had registered on the microblog. We considered two types of social media-related features. We first extracted presence features, including the number of fans, followings, reciprocities, received comments, and "likes" (Ge et al., 2017). Then we calculated the sentiment valence (from -1 = *extremely negative* to 1 = *extremely positive*) and the sentiment variance of each text message the borrower posted on the microblog (Kamath et al., 2013). To identify the sentiment valence, we adopted HowNet sentiment lexicons.[9] We then implemented an in-depth sentence-based sentiment analysis on the microblog messages using the eight lexical categories (denoting sentiments varying from negativity to positivity) predefined in HowNet.[10] The mean sentiment valence of our sample borrowers was 0.02, revealing a neutral attitude. Table 1 summarizes the above features, and Table A1 reports the detailed statistics.

Evaluation Outcomes

Theoretically, a loan becomes delinquent when the borrower makes a payment late, whereas a loan goes into default if the borrower misses several installment payments over a certain period and fails to keep up with ongoing loan obligations. Most financial service providers impose a (relatively large) fine on delinquent borrowers. That is to say, although delinquency incurs a financial loss, from the profit perspective, a delinquent borrower with a certain level of credit risk could be valuable if the borrower repays the installment and fine. Hence, we argue that a comprehensive credit risk assessment should consider not only default probabilities (to avoid high credit risk) but also the probabilities of "delinquent but not in default" (to increase potential revenues). The risk antecedents may likewise have different effects in predicting delinquency and default behavior (Chehrazi & Weber, 2015). Based on the above discussions, we propose three indicators for assessing the credit risk of individuals applying for a microloan.

Delinquent/default: This is a multiclass categorical factor (0 = *not delinquent*, 1 = *delinquent but not in default*, and 2 = *default*). Note that a non-delinquent loan means that there was no delinquent installment repayment across the entire repayment period.

Repayment rate: This variable measures the proportion of repaid monthly installments. Unlike the categorical factor, which captures the overall repayment performance, this numerical indicator delivers finer-grained details on repayment behavior (Drozd & Serrano-Padial, 2017). For example, for a 7-month loan, 0, 4, and 6 repaid installments (i.e., repayment rates equal to 0, 0.57, and 0.86) reflect different levels of risk, though they all belong to the class of "default."

Loan profit: Consistent with prior literature (Papouskova & Hajek, 2019), we calculate the website's profit per loan based on cash flow (or loss given default), by considering both the cost and revenue. The cost includes the loss of principal capital (i.e., exposure at default) and the opportunity cost from a default or delinquent loan. The revenue includes gains of interest and possible penalties (fines) for late payment (delinquency). Our calculation steps in detail are introduced in Appendix B.

Table 2 provides descriptions of the aforementioned three credit risk indicators. Figure B1 (Appendix B) displays the distributions of repayment rates and loan profits. In our sample, 639 borrowers had never repaid their loans. A total of 2,375 loans each yielded positive profits of less than 150 USD, most of which was revenue from interest, and 1,615 (= 817 + 798) loans resulted in financial losses of 300-600 USD. Moreover, the average delinquency duration was approximately 29 days. In sum, Table 2 and Figure B1 demonstrate the high-risk performances of the full applicant pool on the focal microloan website.

[9] http://www.keenage.com/download/sentiment.rar.

[10] The eight lexical categories include positive affectivity (covering 836 words), negative affectivity (covering 1,254 words), and six subcategories of modal intensity (covering 219 words in total): "extreme," "very," "more," "-ish," "insufficiently," and "over."

Table 1. Summary of Constructed Features			
Feature category		**Features #**	**Description/sample features**
Conventional data	Demographic and socioeconomic characteristics, loan attributes, and microloan history	31	- Borrowers' demographic and socioeconomic characteristics - Loan attributes - Loan purpose - Microloan history - Number of defaults of first-order recent cellphone contacts
Alternative data	Online (shopping) activities	40	- Shopping for durable goods - Shopping for virtual products - Consumption-to-income ratio - "Special" product (e.g., alcohol, game card, medicine, book, take-out food, etc.) consumption - Shopping for other people - Transferred in and out on a third-party payment platform
	Mobile activities	31	- Cellphone operation system (iOS or Android) - Cellphone calls and messages - Information on contacts - App usage - Mobility traces
	Social media activities	15	- Whether using social media/microblog - Social media presence - Sentiment

Table 2. Description of Credit Risk Indicators (Evaluation Outcomes)						
Credit risk indicator	**Type**	**Description**	**Mean**	***SD***	**Min**	**Max**
$Delinquent/Default$	Categorical	Multiclass variable. (1 = *not delinquent*, 2 = *delinquent but not in default*, and 3 = *in default*)	784 (15.04%) for class 1, 1,329 (25.45%) for class 2, and 3,101 (59.47%) for class 3			
$Repayment\ rate$	Numerical	# of paid installments/loan period	0.606	0.368	0	1
$Loan\ Profit$	Numerical	Profit of each loan (k, USD)	-0.109	0.328	-1.101	4.222

Model Training Process

Inspired by the industry practice whereby microfinance institutes lend to select applicants according to the anticipated credit risk, we first implemented extreme gradient boosting (XGBoost)[11] to predict individual applicants' delinquent or default behavior. Many studies (e.g., Munkhdalai et al., 2019) have reported that XGBoost achieved the highest credit risk prediction precision. Based on the predicted behavior, our novel experimental design allowed us to examine every counterfactual, which is every possible selection strategy.

To gauge the predictive power of the different feature categories, we first included each individual feature category (*Fc, Fo, Fm, and Fs*), respectively, and treated conventional features (Fc) as the benchmark. Then, we combined features from different categories. Because only one third of our sample had microblog records, we evaluated the prediction performance using two combinations: combining all features except microblog-related features ($Fc \cup Fo \cup Fm$) for the whole sample, and combining all four categories ($Fc \cup Fo \cup Fm \cup Fs$) for the microblogger subsample. All of the features were normalized to ensure the comparability of results. We randomly partitioned our sample into two parts: approximately two thirds (3,476 loans) served as the training and validation sample and the remaining one third (1,738 loans) served as the testing sample. We considered several commonly adopted metrics to evaluate prediction performance. Specifically, for the multiclass categorical risk indicator ($delinquent/default$), the following three metrics were considered: precision, recall, and F1 score. For the

[11] We implemented diverse widely accepted machine learning models, including logistic and linear regression (L&R), support vector machine (SVM), *k*-nearest neighbor (*k*-NN), multilevel perceptron (MLP), and the two most-updated ensemble methods: random forest (RF) and XGBoost. We observed that among all of the machine learning models we considered, XGBoost showed the best performance consistently across the different metrics and feature sets. Appendix C reports the technical details and evaluation results.

numerical risk indicators (repayment rate and loan profit), we considered three evaluation metrics: mean absolute error, root mean squared error, and R^2. The results using the different evaluation metrics showed consistency (Appendix C).[12]

Profitability Analyses

In this section, we evaluate the selection strategies based on the different feature sets adopted in the above XGBoost-based prediction. We define the total profits as profitability.

In practice, the microloan website makes approval decisions on whether to offer an applicant a loan product or not. The most commonly implemented strategy is evaluating a borrower's default probability. This strategy emphasizes the necessity of excluding applicants posing a high credit risk. We used this default-based strategy as our benchmark. In addition, we proposed three alternative business strategies for applicant selection based on their probabilities of delinquent-but-not-in-default behavior and the predicted values of repayment rate and loan profit.

For each of the above four selection strategies, we first predicted the corresponding probabilities/values using different sets of features. We then ranked all of the applicants (from best to worst) based on our predicted values, assuming that the ranking is the only criterion in the approval decision-making process. To evaluate the performance, we then calculated the actual profits by choosing the top 5%, 10%, 15%, … 100% best loans using different thresholds. In Table 3, we report the profits from the approved loans (in the testing set) based on the predicted values of the different credit risk indicators.

We found that the microloan website can generally achieve the highest profits at a 45% loan approval rate across the different credit risk indicators and feature sets (values in bold in Table 3; the test for difference significance in Appendix C). When the loan approval rate is higher than 65%, the microloan business will become unprofitable. More importantly, we found that loan selection based on mobile activity features (*Fm*) yielded the highest economic gains for the microloan website: approximately 22% ((16.95-13.92)/13.92) more economic gains than selection based on conventional information only, under the optimal 45% approval rate. When we applied all of the feature sets to predict credit risk and make loan approval decisions, we found 28% (= (17.76 - 13.92)/13.92) greater economic gains for the microloan website compared to using conventional features only.

Furthermore, comparisons across different evaluation outcomes produced additional interesting findings. For example, when the loan approval rate was lower than 35%, the loan selection strategies using the loan profit indicator and delinquent-but-not-in-default probabilities yielded higher profits than those based on default probabilities and repayment rates. One potential explanation for this is that the loan profit indicator and delinquent-but-not-in-default loans are most economically valuable strategies, especially when applied to guide loan selection decisions.[13] More interestingly still, we found that if the budget only allowed for less than 15% of approved loans, the loan selection strategy using delinquent-but-not-in-default probabilities with mobile activity features only (*Fm*, 9.07, 6.28, 2.85 thousand USD) yielded even higher profits than the strategy using the commonly adopted default probabilities with all features ($Fc \cup Fo \cup Fm \cup Fs$, 8.89, 6.07, 3.00 thousand USD). As such, our results indicate that microloan websites can leverage the combination of alternative feature sets and personalized credit risk indicators to maximize financial profitability within their budgets.

Financial Equality Analyses

Considering the traditional trade-off between profitability and equality, we next explore the question of whether the involvement of alternative data would bring additional benefits to or sacrifice financial equality.

As discussed in the Introduction, financial inclusion helps at least two groups that have traditionally suffered financial inequality to enjoy greater access to financial services: (1) underserved (e.g., thin-file) users and (2) users who have been traditionally thought of as bad credit risks. Thin-file users are often invisible to financial sectors due to the insufficiency of demographic information and loan histories. Our welfare analysis findings revealed higher profitability with alternative data, thus implying that financial service providers can seek alternative data to reach out to underserved users who do not have traditional data recorded in their files. Therefore, next in this section, we investigate the value of alternative data for mitigating the potential unfairness issue in terms of financial inclusion, an issue that has scarcely been addressed in the literature.

[12] Certain geographic areas might be more likely to produce certain types of alternative data, which might result in the prediction model being biased toward/against some of these areas. Our additional prediction analyses (please refer to Appendix F) suggest that this bias effect is trivial.

[13] Selecting loans directly according to loan profits should theoretically be the ideal (most accurate) approach; however, for loan approval rates between 15% and 35%, we found that delinquent-but-not-in-default probabilities led to higher profits than did the direct loan profit indicator, which showed lower accuracy in the prediction analysis.

Table 3. Profits of Alternative Data (k, USD; Microblogger Subsample)

(a) Based on default prediction														
Feature set	**Approved loans (%)**													
	5%	10%	15%	20%	25%	30%	35%	40%	**45%**	50%	55%	60%	65%	70%
Fc	1.65	3.67	5.83	7.48	9.60	10.47	12.42	12.99	13.92	**14.73**	12.53	8.52	1.96	-5.65
Fo	2.26	4.86	7.35	9.69	11.53	12.63	14.35	15.12	**15.75**	15.15	14.26	9.07	1.99	-6.01
Fm	2.47	5.44	8.05	10.44	12.43	13.65	15.33	16.23	**16.96**	16.90	14.74	9.27	1.92	-6.87
Fs	2.46	5.41	7.96	10.32	12.30	13.39	15.01	16.00	**16.45**	16.20	14.70	9.34	1.95	-6.49
Fc∪Fo∪Fm	2.86	5.86	8.67	11.28	13.17	14.58	15.97	17.26	**17.40**	17.17	14.62	9.19	1.77	-7.11
Fc∪Fo∪Fm∪Fs	3.00	6.07	8.89	11.43	13.30	14.91	16.33	17.46	**17.76**	17.20	14.59	9.04	1.74	-7.30
(b) Based on delinquent-but-not-in-default prediction														
Feature set	**Approved loans (%)**													
	5%	10%	15%	20%	25%	30%	35%	40%	**45%**	50%	55%	60%	65%	70%
Fc	1.89	4.21	6.51	8.02	10.03	10.87	12.54	12.75	13.54	**14.31**	13.18	8.61	2.44	-5.01
Fo	2.56	5.49	8.40	10.32	12.13	12.94	14.43	14.86	15.36	**15.39**	13.95	9.30	2.37	-5.35
Fm	2.85	6.28	9.07	11.10	12.90	14.05	15.37	15.90	**16.50**	16.29	14.35	9.25	2.31	-6.15
Fs	2.77	6.16	9.00	10.98	12.70	13.80	15.06	15.58	**15.97**	16.21	14.25	9.46	2.41	-5.97
Fc∪Fo∪Fm	3.43	6.67	9.73	12.04	13.78	15.10	15.99	16.90	**16.87**	16.51	14.31	9.27	2.22	-6.37
Fc∪Fo∪Fm∪Fs	3.51	6.94	10.00	12.22	13.87	15.39	16.44	17.05	**17.31**	16.63	14.23	9.12	2.05	-6.66
(c) Based on Repayment Rate Prediction														
Feature set	**Approved loans (%)**													
	5%	10%	15%	20%	25%	30%	35%	40%	**45%**	50%	55%	60%	65%	70%
Fc	1.63	3.66	5.76	7.54	9.60	10.36	12.51	13.12	**13.98**	13.80	13.20	8.26	1.48	-5.98
Fo	2.28	4.84	7.15	9.70	11.46	12.48	14.43	15.31	15.91	**16.06**	13.95	8.70	1.51	-6.30
Fm	2.43	5.40	8.05	10.30	12.34	13.57	15.49	16.42	**17.07**	16.81	14.44	8.76	1.50	-7.15
Fs	2.38	5.29	7.98	10.17	12.15	13.32	15.10	16.11	**16.63**	15.78	14.29	8.95	1.57	-6.96
Fc∪Fo∪Fm	2.88	5.83	8.67	11.26	13.09	14.52	16.14	17.35	**17.41**	17.14	14.17	8.73	1.35	-7.53
Fc∪Fo∪Fm∪Fs	2.98	6.00	8.77	11.32	13.18	14.82	16.36	17.52	**17.85**	17.14	14.19	8.61	1.27	-7.75
(d) Based on Loan Profit Prediction														
Feature set	**Approved loans (%)**													
	5%	10%	15%	20%	25%	30%	35%	40%	**45%**	50%	55%	60%	65%	70%
Fc	2.14	4.41	6.52	7.89	9.90	10.56	12.30	12.67	**13.68**	13.50	13.47	8.79	2.35	-5.43
Fo	2.94	5.82	8.25	10.21	11.88	12.73	14.20	14.76	**15.46**	14.91	14.19	9.37	2.38	-5.77
Fm	3.22	6.58	9.04	10.98	12.81	13.68	14.98	15.90	16.65	**16.66**	14.67	9.76	2.26	-6.36
Fs	3.19	6.52	8.92	10.87	12.67	13.53	14.77	15.66	**16.15**	15.65	14.62	9.64	2.34	-6.22
Fc∪Fo∪Fm	3.73	7.02	9.70	11.89	13.57	14.71	15.82	16.84	**17.08**	16.90	14.55	9.48	2.11	-6.82
Fc∪Fo∪Fm∪Fs	3.90	7.27	9.97	12.04	13.71	15.04	16.18	17.04	**17.44**	16.95	14.52	9.34	2.08	-7.02

Note: Values in bold indicate the highest profits for the various loan approval rates.

Table 4. Comparison of Top 45% Best Approved Loans (Based on Microblogger Subsample, XGBoost, Default-based Prediction)

		Loan selection strategy				
		Group 1	***Group 2***	***Group 3***	***Group 4***	***Group 5***
		Fc	**Fo**	**Fm**	**Fs**	**Fc∪Fo∪Fm∪Fs**
# (ratio) of overlap loans to Group 5		170 (69.96%)	180 (74.07%)	207 (85.19%)	200 (82.30%)	243 (100%)
Means of samples of conventional features of unique borrowers	City DPI	7,805.98	7,650.42	6,833.21	6,972.04	6,381.09
	Monthly income level	5.28	5.20	4.55	4.69	4.35
	Loan-to-income ratio	1.18	1.20	1.33	1.31	1.40
	Education level	4.10	4.05	3.88	3.90	3.73
	Homeownership	0.50	0.49	0.42	0.42	0.34

Note: All results are based on the full sample-based predictions. Sample features are those showing significantly different mean values across groups.

Based on the experimental data (i.e., "full sample-based prediction" hereafter), by examining and comparing the characteristics of the approved borrowers selected by the strategies with different feature sets, the value of alternative data for enhancing financial inclusion can be explored. Table 4 reports the mean values of the major demographic and socioeconomic features of the approved loan applicants in the five testing sets (i.e., the predictions with different feature sets). Among the five groups, Groups 1 to 4 covered each feature set separately, while Group 5 considered a combination of all four feature sets. A total of 159 loans were selected using all five selection strategies. We learned that Group 1 tended to select applicants with the highest incomes, those from the highest DPI cities, and those with the highest education levels (among the five groups), whereas Group 5 tended to grant loans to a broader population, including applicants with the lowest incomes, those from the lowest DPI cities, and those with the education levels. That is, combining all of the alternative features allowed the website to reach out to more users with less favorable backgrounds. Interestingly, we also observed that Group 5 generated the highest profits among all groups. This finding indicates that with proper design, financial inclusion and companies' economic incentives can be aligned. Our study thus demonstrates the tremendous potential of leveraging alternative data to alleviate inequality in the financial service market while simultaneously realizing greater revenues.

Looking at the loan selection strategy based on each alternative feature set, we found that the selection strategy using mobile activities (Group 3) had the greatest power to improve financial inclusion, followed by the strategy using social media activity information (Group 4). In Table 3, we reveal the value of these two kinds of features for improving profits. These findings suggest that alternative data from smartphone usage records or social media have the potential to contribute to balancing the trade-off between financial profitability and equality. By contrast, using online activities (Group 2) results in approved applicants who have very similar characteristics to those approved through the strategy using conventional features. That is, although online shopping activity features help boost profitability, they have insufficient value for mitigating financial inequality in this context. We also applied the fairness criterion of "equalized opportunity"—namely, that positive outcomes should be independent of the protected attribute (Teodorescu et al., 2021)—to examine whether the full sample and alternative feature sets can promote fairness. We obtained consistent results, which are presented in Appendix G.

Mechanism Detection

In this section, we empirically disentangle potential explanations for why alternative data could help increase both profitability and financial equality. In particular, we propose that alternative data could help reduce training sample biases and empirically test this proposal. We first reveal the existence of training sample biases and then quantify the degree to which alternative data could reduce the biases and realize the corresponding benefits in profitability and financial equality. Next, we take a further step to examine why alternative data could reduce training sample biases using feature-wise correlation analysis.

Alternative Data Help Reduce Training Sample Bias

We first examine why alternative data have merit for improving profitability. As discussed above, prior studies and current industry practice on credit risk prediction tend to rely on approved samples and conventional information. Such data are easily accessible and cleaner, compared with a full applicant pool or richer information. As discussed in the literature review section, however, a lack of comprehensive information for model training and training sample bias are the two main biases that impair the accuracy of credit scoring. In this section, we show that in our setting, a comprehensive feature set extracted from alternative data can not only improve performance but also offset the losses of a biased training sample.

We constructed a counterfactual "approved sample" from our initial training experimental data (with 3,476 loan samples; see the Model Training Process section above). We proceeded as follows: We first divided the 3,476 samples into two parts: a training and validation subsample (1,158) and a testing subsample (2,318).[14] Then, we trained XGBoost[15] based on the 1,158 training subsample using conventional features to predict the default probability of loan applicants in the prediction subsample. According to the predicted default probability, we chose the top 45% of loan applicants (1,043) from the testing subsample as the counterfactual approved sample. It is worth noting that these 1,043 approved samples are an artificial set representing the baseline—which would have been generated in the traditional way, wherein the microloan company predicts future default probabilities for new applicants using historical observations from previously approved samples. In the following analyses, we compare the

[14] We did this to obtain as large a subsample size as possible for the final approved sample (to be used in next-step model training).

[15] We tested with different state-of-the-art machine learning models, and the results showed consistency.

performance using our full experimental data (3,476 samples) with that from this counterfactual approved sample.

The loan attributes in this approved sample were akin to those in the full sample. However, the approved sample had a larger proportion of male borrowers, and the mean values of their socioeconomic characteristics (city DPI, monthly income level, and education level) were higher. They also performed better in terms of microloan history. Compared with the full sample, the approved sample shows a larger variance in the values of demographic and socioeconomic features (i.e., the distribution of these features was more dispersed). This was due to the fact that a large proportion of borrowers in the full sample had performed "worse" but had more homogeneous demographic and socioeconomic characteristics (refer to Appendix D for details). The approved samples showed better repayment performance. Among the approved loans, 336 (i.e., 32.21%) had no delinquent installments, 315 (i.e., 30.20%) belonged to the delinquent-but-not-in-default class, and the other 392 (i.e., 37.59%) were in default. The average repayment rate of these loans was 0.797 (*SD*: 0.353), and the average loan profit was 11 USD (*SD*: 1.105). Overall, these loans had an explicitly lower risk than the full sample.

We applied "approved sample-based prediction." Specifically, we applied the same training strategies on this approved sample as on the full sample. Then, we implemented the coefficients trained from this approved sample to predict the credit risk of the full sample. This loan prediction operation resembles that applied in real-world practice.

As expected, Table 5 indicates that compared with the full sample-based prediction,[16] the approved sample-based prediction showed worse performance on the testing set. The performance gap was rather significant (approximately 57% - 73%) when we applied conventional features only. The performance gap became significantly smaller when we incorporated alternative data. For example, with mobile activities, the prediction performance gap of the $delinquent/default$ indicator decreased from 57.32% to 21.78%. In addition to the prediction performance, Table 5 suggests that the findings also hold for the analysis on associate profitability. These findings strengthen our argument regarding the effectiveness of alternative information in credit risk assessment.

A similar conclusion can be drawn from the comparison of economic gains (Figure 1). Specifically, under the optimal loan approval rate (i.e., 45%, 243 loans in the testing set), the website would have an opportunity loss of 3,460 USD (i.e., 24.88%) if approved sample-based prediction with conventional features were implemented, whereas the loss would be 2,350 USD (i.e., 13.25%) if alternative data were incorporated into the credit risk prediction. More interestingly, we found further strong evidence that alternative data can *offset* the economic loss caused by training sample bias. In particular, even with a biased training sample, the economic gains of incorporating alternative data (15,410 USD) were much larger than when using an unbiased training sample but with conventional data only (13,920 thousand USD). These findings indicate that alternative data can help shrink the economic losses caused by sample bias and that the economic value exceeds that obtained by using conventional data only, even without sample bias.

Next, we aim to identify whether the reduced training sample bias increases the equality level. Table 6 reports the mean values of the major demographic features summarized from loan applicants approved by the strategy based on the approved sample-based prediction. A total of 120 loans were selected by all ten selection strategies (together with the five groups in Table 4). Again, we conclude that alternative features can improve financial inclusion and offset the unfairness caused by training sample bias. For example, we observed that compared with the Group 1 strategy (full sample-based prediction with conventional features), the Group 8 strategy (approved sample-based prediction with mobile activity information) favored lower-income applicants, less-educated applicants, and/or those from less-developed geographic areas—i.e., historically disadvantaged, largely neglected populations.

Furthermore, we observed significant differences among the three types of alternative features in terms of their utility for balancing the trade-off between financial profitability and equality. Specifically, to quantify financial equality improvement, we used the increasing percentages of the overlap loans (Row 1 in Tables 4 and 6) compared with Group 6 (approved sample-based prediction with Fc). This calculation assumed that Group 5 (full sample-based prediction with all features) is optimal and that the corresponding selection strategy could maximize financial inclusion by covering the most needy applicants. We learned that, among the three types of alternative features, profiling user credit risk using smartphone activities yielded an approximately 23.05% (= 80.25% - 57.20%) improvement in financial inclusion, while social media activities yielded only an 18.11% (= 75.31% - 57.20%) improvement. That is, using mobile activity features for credit risk assessment was about 1.3 times more effective in improving financial inclusion.

[16] Since the size of the approved sample for model training (1,043) was smaller than that for the full sample-based prediction (3,476), we randomly drew 1,043 from the 3,476 samples to replicate the full sample-based prediction. This allowed for ruling out the potential interference caused by training sample size. The randomly drawn "full subsample" had very close distributions among all the features (Table D1, Appendix D) and the prediction outcomes were also quite consistent with the original predictions that were based on the entire sample.

Table 5. Comparison of Prediction Performance and Profit (based on XGBoost)

(a) Prediction Performance												
Feature set	**Delinquent/Default (F1 score)**				**Repayment rate (R^2)**				**Loan profit (R^2)**			
	Approved-sample	***Full-sample***	***Bias (%)***	***P-values on bias***	***Approved - sample***	***Full-sample***	***Bias (%)***	***P-values on bias***	***Approved-sample***	***Full-sample***	***Bias (%)***	***P-values on bias***
Fc	0.169	0.396	57.32	0.002***	0.033	0.120	72.50	< 0.001***	0.020	0.048	58.33	0.008***
Fo	0.357	0.458	22.05	0.071*	0.163	0.214	23.83	0.070*	0.093	0.146	36.30	0.030**
Fm	0.535	0.684	21.78	0.059*	0.555	0.788	29.57	0.057*	0.262	0.384	31.77	0.044**
Fs	0.360	0.566	36.40	0.011**	0.432	0.739	41.54	0.022**	0.186	0.342	45.61	0.012**
Fc∪Fo∪Fm	0.532	0.686	22.45	0.060*	0.545	0.791	31.10	0.054*	0.265	0.389	31.88	0.045**
Fc∪Fo∪Fm∪Fs	0.534	0.688	22.38	0.060*	0.550	0.793	30.64	0.057*	0.266	0.390	31.79	0.046**
(b) Profit												
Feature set	**Delinquent/Default**				**Repayment rate**				**Loan profit**			
	Approved-sample	***Full-sample***	***Bias (%)***	***P-values on bias***	***Approved - sample***	***Full-sample***	***Bias (%)***	***P-values on bias***	***Approved-sample***	***Full-sample***	***Bias (%)***	***P-values on bias***
Fc	7.05	13.54	47.93	0.001***	7.15	13.98	48.86	< 0.001***	7.06	13.68	48.39	< 0.001***
Fo	12.11	15.36	21.16	0.035**	12.48	15.91	21.56	0.033**	12.25	15.46	20.76	0.035**
Fm	13.08	16.50	20.73	0.042**	13.64	17.07	20.09	0.052*	13.30	16.65	20.12	0.055*
Fs	10.64	15.97	33.38	0.006***	11.08	16.63	33.37	0.014**	10.91	16.15	32.45	0.020**
Fc∪Fo∪Fm	13.43	16.87	20.39	0.047**	14.05	17.41	19.30	0.057*	13.85	17.08	18.91	0.086*
Fc∪Fo∪Fm∪Fs	13.99	17.31	19.18	0.053*	14.43	17.85	19.16	0.058*	14.22	17.44	18.46	0.088*

Note: Profit in Panel (b) is calculated based on the predicted repayment performance with a loan approval rate of 45%. Bias = (Full sample-based - Approved sample-based) / Full sample-based. We report the *p*-values of the pair-wise *t*-test on the performance between the approved sample-based and full sample-based predictions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

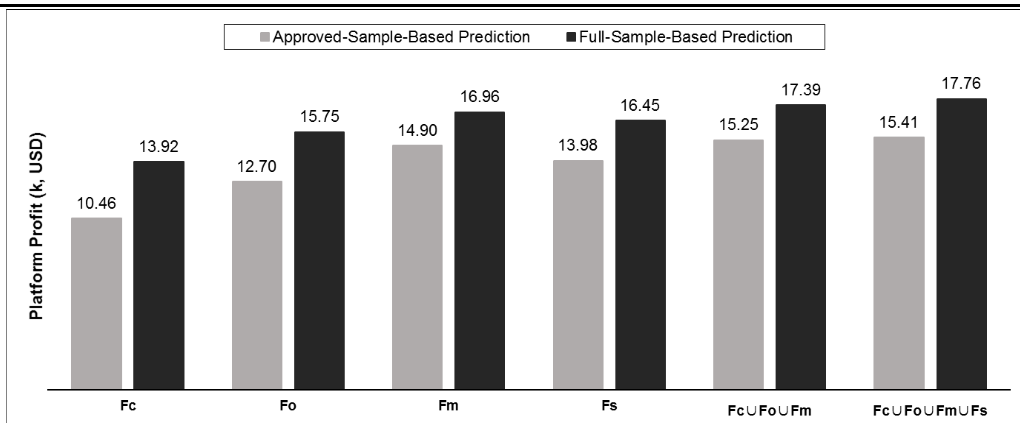


Figure 1. Profit Comparison (based on Microblogger Subsample, XGBoost, Default-based Prediction, Loan Approval Rate 45%)

Table 6. Comparison of Top 45% Best-Approved Loans (based on Microblogger Subsample in Approved Sample, XGBoost, Default-based Prediction)

		Loan selection strategy				
		Group 6	***Group 7***	***Group 8***	***Group 9***	***Group 10***
		Fc	**Fo**	**Fm**	**Fs**	**Fc∪Fo∪Fm∪Fs**
# (ratio) of overlap loans to Group 5		139 (57.20%)	152 (62.55%)	195 (80.25%)	183 (75.31%)	208 (85.60%)
Means of samples of conventional features of unique borrowers	City DPI	8,030.04	7,800.41	7,112.75	7,625.22	6,720.04
	Monthly income level	5.49	5.44	4.77	5.14	4.51
	Loan-to-income ratio	1.06	1.09	1.26	1.22	1.33
	Education level	4.25	4.13	4.00	4.10	3.90
	Homeownership	**0.55**	**0.54**	**0.46**	**0.50**	**0.40**

Note: All results are based on approved sample-based predictions. Sample features are those showing significantly different mean values across groups.

Meanwhile, regarding financial profitability, Figure 1 indicates that profiling user credit risk using smartphone activities yielded a 42% (= (14.90 - 10.46)/10.46) increase in profits; likewise, this was about 1.3 times more effective (an increase of ~33%) in terms of business profitability than using social media information. However, using online shopping activities only (Group 7) did not effect any great reduction (an increase of only 5.35%) in financial inequality, while the increase in financial profitability was approximately 21%. Compared with the highest performance achieved by smartphone activities, online shopping activities presented a larger decrease in the improvement of financial inclusion than they did in the case of profitability (0.19 vs. 0.5 times). That is, our findings show that using strategies based on online shopping activity features would hurt financial inclusion if the financial company applied them for credit risk assessment. This finding is consistent with Tucker (2020), who reported that online cookie-based advertising can hurt poor people's profiles.

In a nutshell, Tables 4 and 6 reveal that adding certain alternative features such as mobile activities and social media activities could help reduce the unfairness caused by training sample bias while simultaneously promoting profitability and financial inclusion. This is a particularly valuable finding, especially given that it is difficult for microloan websites to obtain a full sample in practice.

Effective Alternative Features Are Orthogonal to Conventional Features

The previous section quantified empirically how and to what degree the use of alternative data could help reduce training sample biases. This section further explores why alternative data offer such advantages. Current loan selection strategies are predominantly based on conventional features. Consequently, the training process overemphasizes the "goodness" of conventional features. Unfortunately, many conventional features are based on sensitive demographic attributes such as gender, income, and race. Therefore, loan applicants tend to be discriminated against based on those sensitive attributes, which may cause inadvertent financial inequality.

To verify this, we conducted a further analysis in terms of feature importance. For this purpose, we implemented a permutation feature attribution method (Fisher et al. 2018). Table 7 lists the 10 most important features among the four groups shown in Tables 4 and 6 (Groups 1, 5, 6, and 10). The results explicitly show that conventional features such as city DPI, monthly income level, educational level, and homeownership played significant roles in credit risk assessment for Groups 1 and 6, when only conventional features were employed. By contrast, when alternative features were included (i.e., Groups 5 and 10), these features became more important than the conventional features used in the default-based prediction.

Figure 2 plots the correlations between some important alternative features and conventional sensitive demographic attributes, which suggest that these alternative features do not directly correlate with the sensitive attributes. Moreover, following Thompson (2000), we performed a canonical correlation analysis (CCA) to test the correlation between online shopping activity features and traditional demographic features, as well as the correlation between mobile activity features and traditional demographic features, respectively (refer to Appendices E and H for technical details and results). The results support the assumption that the mobile activity features are orthogonal to the traditional demographic features. The results also support the assumption that the mobile activity features have trivial correlations with the traditional demographic features. This is critical. It suggests that alternative features can better capture borrowers' credit risks—potentially from different angles that are orthogonal to sensitive demographic features such as gender and income. Therefore, a predictive model based on alternative features is less likely to cause bias toward sensitive attributes. Further correlation analyses (see Table E1) suggest that, in general, typical[17] mobile activities and social media information are only weakly correlated with sensitive user attributes such as education level and income level. However, some typical online activities (e.g., purchasing durable goods and virtual products) reveal high correlations with these sensitive attributes. This largely explains why online shipping data cannot effectively contribute to the mitigation of financial inequality.

[17] Typical features here refer to those ranking high in feature-importance analyses.

Table 7. Top 10 Important Features of Different Prediction Strategies (Based on Microblogger Subsample, XGBoost, Default-Based Prediction)

Rank	Group 6 *Approved sample-based prediction with conventional features*	Group 1 *Full sample-based prediction with conventional features*	Group 10 *Approved sample-based prediction with all features*	Group 5 *Full sample-based prediction with all features*
1	City DPI (0.2616)	City DPI (0.2387)	City DPI (0.2007)	# Office by week (0.2107)
2	Monthly income level (0.1869)	Loan amount (0.1597)	Avg_amount_game card (0.1731)	Avg_amount_game card (0.1531)
3	Education level (0.0772)	Monthly income level (0.1361)	# Fans in microblog (0.0835)	# Fans in microblog (0.0775)
4	Loan amount (0.0508)	If_other loan (history) (0.0626)	# Office by week (0.0684)	# Using game apps by week (0.0640)
5	Loan-to-income ratio (0.0446)	# Loans borrowed on microloan websites (0.0381)	# Commercial place by week (0.0604)	# Commercial place by week (0.0574)
6	Homeownership (0.0412)	Gender (0.0344)	Sentiment valence of generated messages (0.0577)	Sentiment valence of generated messages (0.0527)
7	Loan interest rate (0.0362)	Loan-to-income ratio (0.0343)	Monthly income level (0.0570)	# Likes in microblog (0.0510)
8	If_other loan (history) (0.0356)	Education level (0.0325)	# Recreational place by week (0.0511)	# Recreational place by week (0.0509)
9	# Loans borrowed on microloan websites (0.0319)	Loan interest rate (0.0308)	Loan amount (0.0445)	City DPI (0.0385)
10	Loan purpose (0.0314)	Homeownership (0.0300)	# Likes in microblog (0.0379)	Loan amount (0.0259)

Note: The feature-importance score is in parentheses.

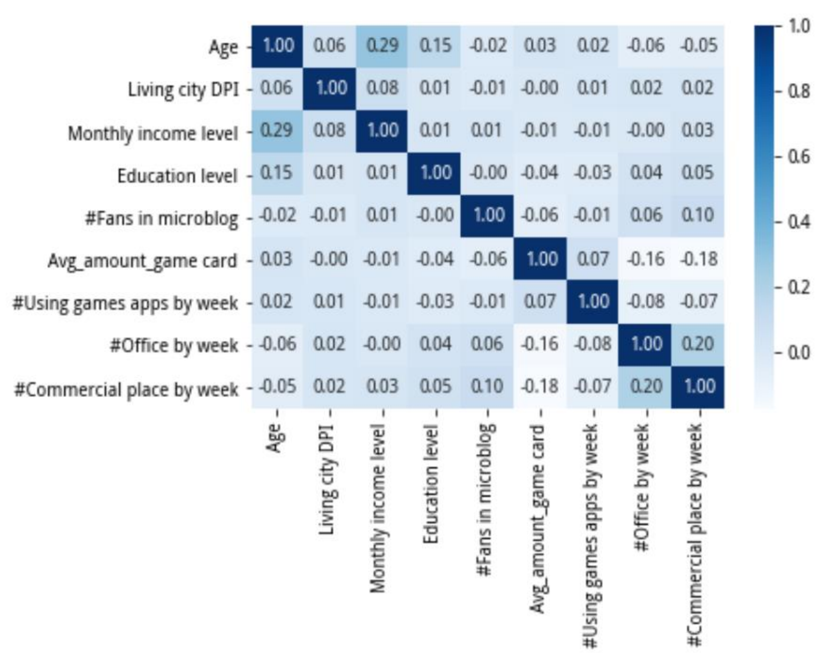


Figure 2. Correlation Matrix of Selected Features (Based on Microblogger Subsample)

Discussion and Conclusion

Summary

By conducting a "meta" experiment to collect unique datasets containing multiple alternative sources of borrower information on a microloan website, we investigated whether alternative data could help balance the trade-off between financial profitability and equality. Our empirical analyses confirmed that with proper selection strategy designs, alternative data increase profits while providing a social lift. Specifically, our findings indicate that using mobile activity data leads to the highest profits. For social media users, social media presence and sentiment are also valuable in enhancing profits. Additionally, both feature sets demonstrate the potential to alleviate financial equality by targeting lower-income applicants, less-educated applicants, or applicants from less-developed areas, who might be neglected in traditional selection strategies. Moreover, our explanation-detection analysis indicates that the mitigation of prediction bias through approved samples and the orthogonality between alternative data and certain sensitive attributes might contribute to the great value of alternative data in balancing the trade-off between financial profitability and equality while enhancing general social welfare. Surprisingly, we found that although using alternative data from online shopping activities can improve business profitability, it does not contribute much to the enhancement of financial inclusion. Indeed, compared with the increasingly popular practice of applying mobile activities to credit risk assessment, using online shopping activities alone threatens the promotion of financial inclusion. In Appendix C, we present a simple back-of-the-envelope calculation of the cost of replicating our study for other websites. In sum, our empirical analyses demonstrate the unique value of alternative data, and such findings have nontrivial implications for other microloan providers regarding the selection and use of different sources of data to promote their ultimate business goals. For example, if a microloan provider aims at enhancing financial inclusion with only a limited budget for collecting/purchasing data, it should consider collecting mobile activities rather than shopping histories.

Data Privacy Issue

We acknowledge that while alternative data could contribute to the realization of both financial profitability and equality, collecting and applying significant amounts of private information could incur operational and legal risks for financial companies (Tang, 2019).[18] To combat this challenge, one possible approach to mitigating the privacy issue of using sensitive private features for microloan websites would be to evaluate borrowers' credit risk with less sensitive metalevel features (i.e., by metafeature-based analysis). Metalevel features are features extracted to describe the distribution of original features corresponding to each sample record. The principal tenet of metafeature-based analysis is to reconstruct borrowers' original feature space in a desensitized way while simultaneously maximizing the retention of discriminative factors embedded in informative features (Ciabattoni et al., 2017; Sharma et al., 2020). Following Rauber et al. (2014), we performed a metafeature-based analysis across different alternative data categories. We applied techniques of feature selection, parameter tuning, model training, and result comparison identical to those employed in the main analyses. The details are reported in Appendix I. Generally, we found that even with the aggregate metalevel features, which were inferior to the original features, alternative data can also help improve the economic gains for the website while simultaneously mitigating financial inequality. Notably, however, directly obtaining users' personal data from third-party sectors and using features extracted from these data might lead to privacy issues. Our findings suggest a feasible alternative approach: user data collection based on the much less sensitive metalevel features between microloan websites and data providers. This would help improve both company and social welfare and would likewise promote information security and the protection of user privacy.

Limitations and Directions for Future Study

Our paper has a few limitations that nonetheless provide promising opportunities for future research. First, since we only had access to individuals' digitized user-behavior data during the loan application stage, we could not dynamically predict their repayment behavior after the approval of their applications. If data are available, future research could extend our analyses to disentangle how and why individuals did not pay off their loans in time. This could facilitate a more efficient and effective assessment of individuals' credit risk. Second, our experimental design was based on fixed interest rates that did not vary with credit risk, the number of

[18] Some regulations have been enacted, such as the General Data Protection Regulation (GDPR) regarding the processing of personal data in the electronic communication sectors.

applications, or any other factors. Although this design would lessen concerns about users' self-section issues, it would affect estimates of profitability because firms are largely free to set interest rates according to individual risk categories. Future studies could relax this condition to explore the economic value of alternative data under varying interest rate schemes based on applicants' financial risk. Third, the issue of cultural background and the scope of the market may have affected our findings. Compared with the growing size of the global microloan market (with overall loans estimated to reach 400 billion USD by 2027), the business scale of the focal microloan website is relatively small. For better generalizability of our findings, future studies could further validate and offer additional insights into the value of information by considering a larger microloan platform in other contexts, especially developed societies with well-established credit systems. Fourth, we performed analyses primarily using a microloan business setting. Given its inherent differences (e.g., money sources and participants) from other settings such as P2P lending, consumer debt, and traditional credit cards, future studies could validate and/or enrich our findings in other interesting financial settings that require financial risk evaluation. For example, individual lenders in P2P lending may be less likely to identify the value of alternative data due to a limited capacity for information processing (Hu et al., 2022). Fifth, our use of data across clusters such as geographic areas, business types, and observation time windows may have resulted in less stable prediction model parameters, potentially causing prediction drift over space and time, which is a caveat to relying on alternative data. Therefore, future studies replicating our analyses in different regions, cultures, and business types could also extend our understanding of the value of alternative data from the methodological perspective. Finally, although we addressed the data privacy issue, we did not dive much into the trade-off between private data collection and borrower participation. This should likewise be related to profitability. Future studies, especially field experiments, could extend the research in this direction.

Acknowledgments

Yingjie Zhang is the corresponding author. We thank the SE, AE, and three anonymous reviewers for their constructive comments. This research was partially supported by the National Natural Science Foundation of China (Grant [72272003] and Grant [71872050]) and a 2019 Facebook Research Award on Economic Opportunity and Digital Platforms.

References

Abbasi, A., Albrecht, C., Vance, A., & Hansen, J. (2012). Metafraud: a meta-learning framework for detecting financial fraud. *MIS Quarterly, 36*(4), 1293-1327. https://doi.org/10.2307/41703508

Amonini, C., & Donovan, R. J. (2005). The relationship between youth's moral and legal perceptions of alcohol, tobacco and marijuana and use of these substances. *Health Education Research, 21*(2), 276-286. https://doi.org/10.1093/her/cyh064

Bassem, B. S. (2012). Social and financial performance of microfinance institutions: is there a trade-off? *Journal of Economics and International Finance, 4*(4), 92-100. https://doi.org/10.5897/jeif11.129

Berentsen, A., & Markheim, M. (2021). Peer-to-peer lending and financial inclusion with altruistic investors. *International Review of Finance, 21*(4), 1407-1418. https://doi.org/10.1111/irfi.12333

Berk, R., Heidari, H., Jabbari, S., Joseph, M., Kearns, M., Morgenstern, J., ... & Roth, A. (2017). *A convex framework for fair regression.* Available at https://arxiv.org/abs/1706.02409

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science, 350*(6264), 1073-1076. https://doi.org/10.1126/science.aac4420

Carroll, P., & Rehmani, S. (2017). *Alternative data and the unbanked.* Oliver Wyman Financial Services. http://www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2017/may/Alternative_Data_And_The_%20Unbanked.pdf.

Chan, T., Hamdi, N., Hui, X., & Jiang, Z. (2020). *Digital verification and inclusive access to credit: An empirical investigation.* Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3556554

Chehrazi, N., & Weber, T. A. (2015). Dynamic valuation of delinquent credit-card accounts. *Management Science, 61*(12), 3077-3096. https://doi.org/10.1287/mnsc.2015.2203

Chouldechova, A. (2017). Fair prediction with disparate impact: a study of bias in recidivism prediction instruments. *Big Data, 5*(2), 153-163. https://doi.org/10.1089/big.2016.0047

Churchill, S. A. (2020). Microfinance financial sustainability and outreach: is there a trade-off? *Empirical Economics, 59*(3), 1329-1350. https://doi.org/10.1007/s00181-019-01709-1

Ciabattoni, L., Ferracuti, F., Freddi, A., & Monteriu, A. (2017). Statistical spectral analysis for fault diagnosis of rotating machines. *IEEE Transactions on Industrial Electronics, 65*(5), 4301-4310. https://doi.org/10.1109/TIE.2017.2762623

Cowgill, B. (2018). *The econometrics of gatekeeping experiments* [Working paper].

Cowgill, B. (2019). *Bias and productivity in humans and machines.* Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3584916

Cowgill, B., & Tucker, C. E. (2019). *Economics, fairness and algorithmic bias.* Available at http://conference.nber.org/confer/2019/YSAIf19/SSRN-id3361280.pdf

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. In *Proceedings of the 21st ACM Conference on Economics and Computation* (pp. 679-681). https://doi.org/10.1145/3391403.3399545

Cui, R., Gallino, S., Moreno, A., & Zhang, D. J. (2018). The operational value of social media information. *Production and*

Operations Management, 27(10), 1749-1769. https://doi.org/10.1111/poms.12707

Cull, R., Demirguç-Kunt, A., & Morduch, J. (2011). Microfinance trade-offs: regulation, competition and financing. In B. Armendáriz & M. Labie (Eds.), *The Handbook of Microfinance*, pp. 141-157. https://doi.org/10.1142/9789814295666_0007

Davis C. (2021). Driving purpose and profit through financial inclusion: stronger together. https://www2.deloitte.com/us/en/insights/industry/financial-services/purpose-through-inclusive-finance.html

Demirguc-Kunt, A., Klapper, L., & Singer, D. (2017). *Financial inclusion and inclusive growth: a review of recent empirical evidence.* Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2958542

Dev, S. M. (2006). Financial inclusion: Issues and challenges. *Economic and Political Weekly, 41*, 4310-4313. https://www.jstor.org/stable/4418799

Dobbie, W., Liberman, A., Paravisini, D., & Pathania, V. (2021). Measuring bias in consumer lending. *The Review of Economic Studies, 88*(6), 2799-2832. https://doi.org/10.1093/restud/rdaa078

Drozd, L. A., & Serrano-Padial, R. (2017). Modeling the revolving revolution: the debt collection channel. *The American Economic Review, 107*(3), 897-930. https://doi.org/10.1257/aer.20131029

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research, 20*(177), 1-81. https://jmlr.org/papers/v20/18-760.html

Fu, R., Aseri, M., Singh, P. V., & Srinivasan, K. (2022). “Un” fair machine learning algorithms. *Management Science, 68*(6), 4173-4195. https://doi.org/10.1287/mnsc.2021.4065

Fu, R., Huang, Y., & Singh, P. V. (2021). Crowds, lending, machine, and bias. *Information Systems Research, 32*(1), 72-92. https://doi.org/10.1287/isre.2020.0990

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance, 77*(1), 5-47. https://doi.org/10.1111/jofi.13090

Ge, R., Feng, J., Gu, B., & Zhang, P. (2017). Predicting and deterring default with social media information in peer-to-peer lending. *Journal of Management Information Systems, 34*(2), 401-424. https://doi.org/10.1080/07421222.2017.1334472

Guild, J. (2017). Fintech and the future of finance. *Asian Journal of Public Affairs*, *10*(1), Article e4.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of The Econometric Society*, *47*(1), 153-161. https://doi.org/10.2307/1912352

Heckman, J.J. & Vytlacil, E.J. (2007). Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. *Handbook of Econometrics, 6*(b), 4875-5143. https://doi.org/10.1016/S1573-4412(07)06071-0

Hu, X., Huang, Y., Li, B., & Lu, T. (2022). *Human-algorithmic bias: Source, evolution, and impact.* Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4195014

Hua, J., Xiong, Z., Lowey, J., Suh, E., & Dougherty, E.R., (2005). Optimal number of features as a function of sample size for various classification rules. *Bioinformatics, 21*(8), 1509-1515. https://doi.org/10.1093/bioinformatics/bti171

Iyer, R., Khwaja, A. I., Luttmer, E. F., & Shue, K. (2016). Screening peers softly: inferring the quality of small borrowers. *Management Science, 62*(6), 1554-1577. https://doi.org/10.1287/mnsc.2015.2181

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

Kamath, S. S., Bagalkotkar, A., Kandelwal, A., Pandey, S., & Poornima, K. (2013). Sentiment analysis based approaches for understanding user context in web content. In *Proceedings of the 2013 International Conference on Communication Systems and Network Technologies* (pp. 607-611). https://doi.org/10.1109/CSNT.2013.130

Kim, E. J., Namkoong, K., Ku, T., & Kim, S. J. (2008). The relationship between online game addiction and aggression, self-control and narcissistic personality traits. *European Psychiatry, 23*(3), 212-218. https://doi.org/10.1016/j.eurpsy.2007.10.010

Kim, Y., Briley, D. A., & Ocepek, M. G. (2015). Differential innovation of smartphone and application use by sociodemographics and personality. *Computers in Human Behavior, 44,* 141-147. https://doi.org/10.1016/j.chb.2014.11.059

Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: friendship networks and information asymmetry in online peer-to-peer lending. *Management Science, 59*(1), 17-35. https://doi.org/10.1287/mnsc.1120.1560

Liu, D., Brass, D. J., Lu, Y., & Chen, D. (2015). Friendships in online peer-to-peer lending: pipes, prisms, and relational herding. *MIS Quarterly*, *39*(3), 729-742. https://www.jstor.org/stable/26629629

Loufield, E., Ferenzy, D., & Johnson, T. (2018). *Accelerating financial inclusion with new data.* Available at http://www.centerforfinancialinclusion.org/storage/AcceleratingFINewData_Final_2018,6

Lu, T., Xu, Y. J., Chen, G., & Zhang, C. H. (2020). *Your posts expose your personality: The value of social media information to credit risk prediction in microloan.* Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4138565

Luchetti, M., Sutin, A. R., Delitala, A., Stephan, Y., Fiorillo, E., Marongiu, M., ..., & Terracciano, A. (2018). Personality traits and facets linked with self-reported alcohol consumption and biomarkers of liver health. *Addictive Behaviors, 82,* 135-141. https://doi.org/10.1016/j.addbeh.2018.02.034

Ma, L., Zhao, X., Zhou, Z., & Liu, Y. (2018). A new aspect on P2P online lending default prediction using meta-level phone usage data in China. *Decision Support Systems, 111,* 60-71. https://doi.org/10.1016/j.dss.2018.05.001

Maskara, P. K., Kuvvet, E., & Chen, G. (2021). The role of P2P platforms in enhancing financial inclusion in the United States: An analysis of peer-to-peer lending across the rural-urban divide. *Financial Management*. 50(3), 747-774. https://doi.org/10.1111/fima.12341

Mehrotra, R., Bhattacharya, P., Tan, T., & Phan, T. (2017). Predictive power of online and offline behavior sequences: evidence from a micro-finance context. In *Proceedings of the International*

Conference on Information Systems. https://aisel.aisnet.org/icis2017/DataScience/Presentations/28

Meyer, H., & Pebesma, E., (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models, *Methods in Ecology and Evolution, 12*(9), 1620-1633. https://doi.org/10.1111/2041-210X.13650

Meyer, H., Reudenbach, C., Ludwig, M., Nauss, T., & Meyer, M. H. (2020). *Package "CAST."* CRAN.

Munkhdalai, L., Munkhdalai, T., Namsrai, O. E., Lee, J. Y., & Ryu, K. H. (2019). An empirical comparison of machine-learning methods on bank client credit assessments. *Sustainability*, *11*(3), 699-721. https://doi.org/10.3390/su11030699

Onnela, J. P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertesz, J., & Barabási, A. L. (2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences, 104*(18), 7332-7336. https://doi.org/10.1073/pnas.0610245104

Óskarsdóttir, M., Bravo, C., Sarraute, C., Vanthienen, J., & Baesens, B. (2019). The value of big data for credit scoring: enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing, 74,* 26-39. https://doi.org/10.1016/j.asoc.2018.10.004

Papouskova, M., & Hajek, P. (2019). Two-stage consumer credit risk modelling using heterogeneous ensemble learning. *Decision Support Systems, 118,* 33-45. https://doi.org/10.1016/j.dss.2019.01.002

Rauber, T. W., de Assis Boldt, F., & Varejão, F. M. (2014). Heterogeneous feature models and feature selection applied to bearing fault diagnosis. *IEEE Transactions on Industrial Electronics, 62*(1), 637-646. https://doi.org/10.1109/TIE.2014.2327589

Roberts, J., Yaya, L., & Manolis, C. (2014). The invisible addiction: cell-phone activities and addiction among male and female college students. *Journal of Behavioral Addictions, 3*(4), 254-265. https://doi.org/10.1556/jba.3.2014.015

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers, 3*(2), 35-146. https://www.jstor.org/stable/2662082

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*(5), 688-701. https ://psycnet.apa.org/record/1975-06502-001

Ruiz, S., Gomes, P., Rodrigues, L., & Gama, J. (2017). Credit scoring in microfinance using non-traditional data. In *Proceedings of the EPIA Conference on Artificial Intelligence* (pp. 447-458). Springer. https://doi.org/10.1007/978-3-319-65340-2_37

Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. *Decision Support Systems, 89,* 113-122. https://doi.org/10.1016/j.dss.2016.06.014

Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. *PloS One, 10*(10), Article e0139427. https://doi.org/10.1371/journal.pone.0139427

Sethi, D., & Acharya, D. (2018). Financial inclusion and economic growth linkage: some cross country evidence. *Journal of Financial Economic Policy, 10*(3), 369-385. https://doi.org/10.1108/JFEP-11-2016-0073

Sharma, G., Umapathy, K., & Krishnan, S. (2020). Trends in audio signal feature extraction methods. *Applied Acoustics, 158,* Article 107020. https://doi.org/10.1016/j.apacoust.2019.107020

Soto, V., Frias-Martinez, V., Virseda, J., & Frias-Martinez, E. (2011). Prediction of socioeconomic levels using cell phone records. In *Proceedings of the International Conference on User Modeling, Adaptation, and Personalization* (pp. 377-388). Springer. https://doi.org/10.1007/978-3-642-22362-4_35

Tan, T., Bhattacharya, P., & Phan, T. (2016). Credit-worthiness prediction in microfinance using mobile data: a spatio-network approach. In *Proceedings of International Conference on Information Systems*. https://aisel.aisnet.org/icis2016/Ebusiness/Presentations/28

Tang, H., (2019). *The value of privacy: evidence from online borrowers.* Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3880119

Teodorescu, M. H., Morse, L., Awwad, Y., & Kane, G. C. (2021). Failures of fairness in automation require a deeper understanding of human-ML augmentation. *MIS Quarterly*, *45*(3), 1483-1499. https://doi.org/10.25300/MISQ/2021/16535

Thompson, B. (2000). Canonical correlation analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics* (pp. 285–316). American Psychological Association.

Tucker, C. E. (2020). *Does black-box profiling lead to disadvantages for less privileged and unhealthy people* [Working paper].

Wang, T., Lu, G., & Yan, P. (2019). A novel statistical time-frequency analysis for rotating machine condition monitoring, *IEEE Transactions on Industrial Electronics, 67*(1), 531-541. https://doi.org/10.1109/TIE.2019.2896109

Wei, Y., Yildirim, P., Van den Bulte, C., & Dellarocas, C. (2015). Credit scoring with social network data. *Marketing Science, 35*(2), 234-258. https://doi.org/10.1287/mksc.2015.0949

Yawe, B., & Prabhu, J. (2015). Innovation and financial inclusion: a review of the literature. *Journal of Payments Strategy & Systems, 9*(3), 215-228. https://www.ingentaconnect.com/content/hsp/jpss/2015/00000009/00000003/art00002

Zhang, Y., Nam, C. S., Zhou, G., Jin, J., Wang, X., & Cichocki, A. (2018). Temporally constrained sparse group spatial patterns for motor imagery BCI. *IEEE Transactions on Cybernetics, 49*(9), 3322-3332. https://doi.org/10.1109/TCYB.2018.2841847

About the Authors

Tian Lu is an assistant professor of Information Systems at W. P. Carey School of Business, Arizona State University. Before joining ASU, he was a postdoc researcher at Carnegie Mellon University. He received his Ph.D. degree in information systems management from Fudan University. His research interests include fintech, the impact of IT, human-AI interfaces, and the sharing economy. His research work has been published in the *Management Science*, *MIS Quarterly*, *Journal of Association for Information Systems*, and *Production and Operations Management*. He has received Best Paper awards at several premier conferences including the International Conference on Information Systems (ICIS 2019), China Summer Workshop on Information Management (CSWIM 2019, 2021), and the Pacific Asia Conference on Information Systems (PACIS 2017).

Yingjie Zhang is an associate professor of marketing at Guanghua School of Management, Peking University. Before joining PKU, she was an assistant professor of information systems at the University of Texas at Dallas. She received B.S. degrees in computer science and economics from Tsinghua University and a

Ph.D. degree in information systems and management from the Heinz College, Carnegie Mellon University. Her research interests center on fintech, mobile and sensor technologies, big data and smart cities, user-generated content, and the sharing economy. Her research work has been published in the *Information Systems Research*, *Transportation Research Part C*, and *ACM Transactions on Intelligent Systems and Technology*. She received a Best Paper award at the International Conference on Information Systems (ICIS 2019) and the INFORMS ISS Nunamaker-Chen Dissertation Award (2019). She recently was the runner-up for the Best Published Paper Award for *Information Systems Research* in 2020.

Beibei Li is the Anna Loomis McCandless Chair and Professor of IT & Management at the H. John Heinz III College of Carnegie Mellon University. She has extensive experience leveraging large-scale observational data analytics and experimental analysis with a strong focus on modeling individual user behavior across online, offline, and mobile channels for decision support. She is the winner of several Best Paper awards and is the recipient of over $3M in research awards and grants. Beibei received the INFORMS ISS Sandy Slaughter Early Career Award. She is also the recipient of the INFORMS ISS Nunamaker-Chen Dissertation Award and the ACM SIGMIS Best Doctoral Dissertation Award.

Appendix A

Feature Descriptions

Table A1. Description of Constructed Features

(a) Conventional features (Fc)						
Feature		**Description**	**Mean**	***SD***	**Min**	**Max**
(1)	Gender	Binary; 1 = *female*, 0 = *male*	0.20	0.40	0	1
(2)	Age	–	24.36	3.45	19	46
(3)	City DPI	Disposable personal income (DPI; in USD) in 2017 of residents in the city a borrower lives in	5,055	2,339	3,242	13,510
(4)	Monthly income level	1 = *150 USD or less*: 620 (11.89%); 2 = *150-300 USD*: 1,001 (19.20%); 3 = *300-450 USD*: 1,441 (27.64%); 4 = *450-600 USD*: 1,257 (24.11%); 5 = *600-750 USD*: 655 (12.56%); 6 = *750-900 USD*: 193 (3.70%); 7 = *900-1,050 USD*: 35 (0.67%); 8 = *1,050-1,200 USD*: 10 (0.19%); 9 = *1,200 USD or above*: 2 (0.04%).				
(5)	Loan-to-income ratio	= (11)/((4) × 1,000 - 500)	1.62	1.40	0.15	13
(6)	Homeownership	Binary; 1 = *self-owned*, 0 = *other*.	0.18	0.38	0	1
(7)	Marital status	Binary; 1 = *married*, 0 = *other*.	0.10	0.30	0	1
(8)	# Children	Number of children	0.10	0.34	0	6
(9)	Education level	1 = *middle school or less*: 54 (1.04%); 2 = *vocational school*: 572 (10.97%); 3 = *high school*: 2,237 (42.90%); 4 = *technical school*: 2,016 (38.67%); 5 = *undergraduate*: 328 (6.29%); 6 = *postgraduate*: 7 (0.13%).				
(10)	# Registered contacts	# Registered persons with a close relationship to borrowers	2.42	0.85	1	4
(11)	Loan amount	Loan size (in USD)	465	82	75	1,350
(12)	Loan term	Loan period in months	5.65	1.74	1	7
(13)	Loan interest rate	Yearly loan interest rate (%)	13.70	1.43	12	16
(14)	Loan purpose	Binary; 1 = *for (high) consumption (e.g., traveling)*, 0 = *for dealing with emergencies (e.g., healthcare, accidents, or business turnover)*	0.50	0.50	0	1
(15)	Due_date_Holiday	Binary; 1= *installment due date was during holidays,* 0 = *not*	0.10	0.30	0	1
(16)	Due_date_Weekends	Binary; 1= *installment due date is during weekends,* 0 = *not*	0.28	0.45	0	1
(17)	Due_date_B/E Months	Binary; 1= *installment due date was during the beginning or end (3 days) of a month*, 0 = *not*	0.20	0.40	0	1
(18)	If_other loan (history)	Binary; whether a borrower borrowed money from the focal microloan website	0.34	0.47	0	1
(19)	# Contacted microloan websites	Number of microloan websites a borrower contacted with	1.40	1.88	0	25
(20)	# Contacts with microloan websites	= (21) + (22)	4.16	8.54	0	190
(21)	# Call out to microloan websites	Times a borrower called out to microloan websites in the country	2.68	5.84	0	190
(22)	# Call in from microloan websites	Times calling in from microloan websites in the country to a borrower	1.48	2.33	0	65
(23)	# Registered microloan websites	# Accounts registered on microloan websites in the country	8.28	2.10	0	26
(24)	If_other loan (current)	Binary; 1 = *a borrower had debt on the cooperated websites of the focal microloan website*, 0 = *not*	0.002	0.04	0	1
(25)	# Loans borrowed on microloan websites	Times a borrower borrowed money from microloan websites in the country	0.09	0.46	0	5
(26)	If_delinquent on other microloan websites	Binary; 1 = *a borrower once became delinquent on other microloan websites*, 0 = *not*	0.36	0.48	0	1

(27)	# Defaults_first-order contacts	# Loan defaults a borrower's first-order contacted persons (in recent cellphone call logs) had in the focal and other microloan websites	0.07	0.43	0	6
(28)	If_credit card	Binary; 1 = *a borrower owns credit card(s)*, 0 = *not*	0.05	0.22	0	1
(29)	If_pay credit card regularly	Binary; 1 = *a borrower pays their credit card regularly every month*, 0 = *not*	0.03	0.17	0	1
(30)	Type of occupation	1 = *self-employed*: 1,821 (34.93%); 2 = *private enterprise*: 2,586 (49.60%); 3 = *foreign company*: 107 (2.05%); 4 = *state-owned enterprise*: 594 (11.39%); 5 = *government or public institution*: 106 (2.03%).				
(31)	If_insurance	Binary; 1 = *a borrower has insurance (endowment, medical, unemployment, etc.)*, 0 = *not*	0.54	0.50	0	1
		(b) Online (Shopping) Activity Features (Fo)				
Feature		**Description**	**Mean**	***SD***	**Min**	**Max**
(32)	Amount_transfer out	Total amount a borrower transferred out on a widely accepted third-party payment platform	451	1,806	0	9,630
(33)	Amount_spent	Total amount a borrower spent on a widely accepted third-party payment platform	1,050	2,849	0	95,730
(34)	Amount_transfer in	Total amount a borrower transferred in through a widely accepted third-party payment platform	848	1,993	0	82,370
(35)	Ratio_transfer out-to-in	= (32) / (34)	0.54	4.13	0	508
(36)	If_virtual credit	Binary; 1 = a borrower has virtual credit service account on a widely accepted third-party payment platform, 0 =not	0.76	0.43	0	1
(37)	Avg_amount_takeout food	Average amount of ordering take-out food	92.30	48.21	0	235
(38)	Avg_amount_game card	Average amount of game card purchase or top-up	9.02	12.63	0	118
(39)	Avg_amount_phone top-up	Average amount of cellphone top-up	6.16	5.40	0	36
(40)	Amount_shopping_durable	Total amount of durable product consumption	359	750	0	14,123
(41)	# Order_shopping_durable	Total number (frequency) of durable product consumption orders	20.80	33.14	0	603
(42)	ATV_shopping_durable	= (40) / (41)	115.19	298.45	0	8,000
(43)	# Product_shopping_durable	Total number of purchased durable products	80.84	932.17	0	32,035
(44)	Diversity_shopping_durable	Total number of purchased durable product categories	3.59	1.84	0	6
(45)	Amount_shopping_virtual	Total amount of virtual product consumption	387	930	0	42,915
(46)	# Order_shopping_virtual	Total number (frequency) of virtual product consumption orders	30.12	33.08	0	515
(47)	ATV_shopping_virtual	= (45) / (46)	85.76	159.87	0	6,008
(48)	# Product_shopping_virtual	Total number of purchased virtual products	220.40	3,951	0	286,100
(49)	Ratio_shopping_amount-to-income	= ((40) + (45)) / ((4) × 1,000 - 500)	1.81	0.98	0.010	43
(50)	Variance_amount_shopping	Standard deviation of weekly amount of durable product consumption	122.63	250.20	0	12,152
(51)	Variance_# order_shopping	Standard deviation of the weekly number of durable product consumption orders	4.88	4.57	0	130.26
(52)	# Order_alcohol	Total amount of purchased alcohol	0.63	1.80	0	55
(53)	# Order_caffeine	Total amount of purchased caffeine	0.16	0.59	0	15
(54)	# Order_tobacco	Total amount of purchased tobacco	0.002	0.02	0	2
(55)	# Order_book	Total number of purchased books	0.53	2.17	0	138
(56)	# Order_medicine	Total number of purchased medicine/drugs	0.26	0.81	0	20
(57)	# Order_adult products	Total number of purchased adult products	0.08	0.41	0	12
(58)	Amount_alcohol	Total amount (in USD) of purchased alcohol	18.53	102.36	0	3,610
(59)	Amount_caffeine	Total amount (in USD) of purchased caffeine	2.37	23.78	0	1,027
(60)	Amount_tobacco	Total amount (in USD) of purchased tobacco	0.01	0.42	0	24
(61)	Amount_book	Total amount (in USD) of purchased books	6.93	84.06	0	5,293

(62)	Amount_medicine	Total amount of purchased medicine/drugs	3.13	19.76	0	730
(63)	Amount_adult products	Total amount of purchased adult products	0.97	7.22	0	303
(64)	Ratio_alcohol to durable amount	= (58) / (40)	0.05	1.55	0	114.04
(65)	Ratio_caffeine to durable amount	= (59) / (40)	0.006	0.032	0	1
(66)	Ratio_tobacco to durable amount	= (60) / (40)	0.000	0.000	0	0.16
(67)	Ratio_book to durable amount	= (61) / (40)	0.02	0.17	0	9.67
(68)	Ratio_medicine to durable amount	= (62) / (40)	0.009	0.04	0	1
(69)	Ratio_adult products to durable amount	= (63) / (40)	0.003	0.026	0	1
(70)	# Order_shopping_for others	Total number (frequency) of consumption for other people	6.88	14.89	0	395
(71)	Ratio_shopping_for others	= (70) / ((41) + (46))	0.14	0.21	0	1
(c) Mobile Activity Features (Fm)						
Feature		**Description**	**Mean**	***SD***	**Min**	**Max**
(72)	# Calls by month	= (73) + (75)	180.96	170.11	0	1,720
(73)	# Calls out by month	Average monthly number of outgoing calls	86.75	87.45	0	1,136
(74)	# Duration calls out by month	Average monthly duration (mins) of outgoing calls	127.26	155.00	0	3,896
(75)	# Calls in by month	Average monthly number of incoming calls	94.21	86.56	0	890.4
(76)	# Duration calls in by month	Average monthly duration (mins) of incoming calls	117.50	125.18	0	2,046
(77)	Amount_phone extra expense	Average monthly extra cellphone expenses beyond cellphone plan	11.85	40.14	0	2,965
(78)	Ratio_# call in-to-out	= (75) / (73)	1.09	1.77	0	36.8
(79)	Ratio_duration call in-to-out	= (76) / (74)	0.93	3.20	0	210.85
(80)	# Outgoing contacts	Average monthly # outgoing unique contacted persons	6.50	6.62	0	126.7
(81)	# City_outgoing contacts	Average monthly # cities outgoing contacted persons are in	4.49	6.01	0	92.8
(82)	# Incoming contacts	Average monthly # incoming unique contacted persons	6.49	6.41	0	139.8
(83)	# City_incoming contacts	Average monthly # cities incoming contacted persons are in	4.59	4.37	0	70.6
(84)	# SMS received by month	Average monthly # short text messages a borrower received	73.16	86.22	3	1,481
(85)	# SMS sent by month	Average monthly # short text messages a borrower sent	104.41	89.13	7	1,186
(86)	Ratio_SMS_received-to-sent	= (84) / (85)	0.70	0.64	0.02	13.15
(87)	Cellphone system	Cellphone operation system; 0 = iOS, 1 = Android	0.34	0.47	0	1
(88)	Phone number registered duration	Duration (in months) since the cellphone number was registered (started using) by a borrower	39.58	25.86	6	248
(89)	# Day_phone_longest silence	The longest single duration of the cellphone number keeping silent (i.e., no calling or messaging happened) in history	6.30	14.88	0	361
(90)	# Data usage by month	Average monthly data usage	804.55	448.76	4.68	3,509
(91)	# Apps in cellphone	Number of apps installed in a borrower's cellphone	91.10	37.87	14	189
(92)	# Financial apps	Number of financial and payment apps installed in a borrower's cellphone	4.45	2.40	1	12
(93)	# Using financial apps by week	Average weekly times a borrower used financial and payment apps	3.14	2.21	0	21.9
(94)	# Using social media apps by week	Average weekly times a borrower used social media apps	15.50	10.21	2.5	90.8

	Feature	Description	Mean	*SD*	Min	Max
(95)	# Using entertainment apps by week	Average weekly times a borrower used entertainment (e.g., video) apps	7.90	5.31	0	33.2
(96)	# Using game apps by week	Average weekly times a borrower used game apps	7.75	6.60	0	33.8
(97)	# Using news apps by week	Average weekly times a borrower used news apps	8.00	6.78	0	32.0
(98)	# Cities traveled	Total number of cities a borrower appeared in	2.18	1.90	1	18
(99)	# Office by week	Average weekly frequency (times) of appearance in office buildings/areas	15.84	5.72	3.4	34.8
(100)	# Recreational place by week	Average weekly frequency (times) of appearance in entertainment/recreational places (e.g., movie theatres and amusement parks)	1.03	0.83	0	5.1
(101)	# Commercial place by week	Average weekly frequency (times) of appearance in commercial places (e.g., shopping malls and restaurants)	4.06	4.20	0	26.2
(102)	# Public service place by week	Average weekly frequency (times) of appearance in public service places (e.g., schools and hospitals)	3.70	2.83	0	16.4
(d) Social media activity features (Fs)						
Feature		**Description**	**Mean**	***SD***	**Min**	**Max**
(103)	If_microblog	Binary; 1 = a borrower uses microblog, 0 = *n*ot	0.31	0.46	0	1
(104)	# Fans in microblog	Number of fans a borrower has in microblog	73.66	180.23	0	1,715
(105)	# Following in microblog	# Followings a borrower received in microblog	88.12	197.01	0	1,695
(106)	# Reciprocity in microblog	# Reciprocated relationships (mutual fans) a borrower received in microblog	14.55	26.70	0	285
(107)	# Messages in microblog	= (108) + (109)	73.95	160.63	0	1,414
(108)	# Generated messages in microblog	# Originally generated messages in a borrower's microblog	20.14	54.05	0	640
(109)	# Forwarded messages in microblog	# Forwarded messages from others in a borrower's microblog	53.81	117.18	0	1,175
(110)	# Likes in microblog	# "Likes" a borrower received in microblog	38.71	103.66	0	1,128
(111)	# Comments in microblog	# Comments a borrower received in microblog	46.14	112.58	0	1,277
(112)	Sentiment valance	Average sentiment valence of messages in a borrower's microblog; from -1 (extremely negative) to 1 (extremely positive)	0.02	0.12	-0.97	0.98
(113)	Sentiment valance of generated messages	Average sentiment valence of originally generated messages in a borrower's microblog; from -1 (extremely negative) to 1 (extremely positive)	0.02	0.12	-0.97	0.98
(114)	Sentiment valance of forwarded messages	Average sentiment valence of forwarded messages in a borrower's microblog; from -1 (extremely negative) to 1 (extremely positive)	0.02	0.12	-0.96	0.94
(115)	Sentiment variance	Standard deviation of sentiment valence of messages in a borrower's microblog	0.02	0.04	0	0.15
(116)	Sentiment variance of generated messages	Standard deviation of sentiment valence of originally generated messages in a borrower's microblog	0.02	0.04	0	0.15
(117)	Sentiment variance of forwarded messages	Standard deviation of sentiment valence of forwarded messages in a borrower's microblog	0.02	0.04	0	0.16

Note: Binary feature (103) *If_microblog* was used for sample selection of microblog users in the first stage of the Heckman two-stage approach.

Table A2. Comparison of Borrower and Loan Characteristics across Interest Rates									
Interest rate	# Loans (proportion)	Default rate	Borrower characteristics				Loan characteristics		
			Gender (1 = female)	Age	Monthly income level	Education level	Loan amount	Loan term	Loan purpose (1 = consumption)
12%	1,476 (28.31%)	57.79%	0.20	24.24	3.18	3.32	466.65	5.66	0.49
13%	1,126 (21.59%)	60.04%	0.21	24.04	3.22	3.28	460.16	5.67	0.52
14%	860 (16.49%)	59.42%	0.20	24.60	3.17	3.31	468.41	5.62	0.50
15%	980 (18.80%)	60.20%	0.20	24.36	3.24	3.24	465.48	5.65	0.51
16%	772 (14.81%)	61.01%	0.20	24.80	3.19	3.33	466.90	5.66	0.49
p-value for between-group *F*-test	–	0.58	0.92	0.73	0.85	0.33	0.84	0.96	0.49

Example of the Focal Microblog

The focal social medium is the largest microblog community in the country. It is the mirror application of Twitter. Users can post and repost messages and interact with others on their home page at the microblog website. Figure A1 shows an example of a user's home page on the microblog website.

Figure A1. Example of User's Home Page on Microblog Website

Appendix B

Loan Profit Calculation

Step 1: Revenue Calculation

On the focal website, borrowers are requested to pay their loans every month (i.e., in installments) in equal amounts and with interest. The revenue comes from non-delinquent and delinquent-but-not-in-default loans. We first defined loan i's capital repayment at installment t as $A_{it} = \frac{A_i}{T_i}$, where A_i is loan i's total amount (i.e., principal capital) and T_i is the corresponding period, $t \in (1,2, \dots, T_i)$. We then defined the per-installment interest rate as: $IR_{it} = \frac{IR_i}{12}$, where IR_{it} is loan i's yearly interest rate. The due date of loan i's installment t is denoted as D_{it}^d and the actual payment date is denoted as D_{it}^p. We also defined:

$$y_{it} = \begin{cases} 0, & loan\ i's\ \ installment\ t\ was\ not\ paid; \\ 1, & loan\ i's\ installment\ t\ was\ paid. \end{cases}$$

Specifically, there are four cases of revenue calculation. **Case 1:** An installment is paid on time (i.e., $D_{it}^d = D_{it}^p$). Loan interest is the only source of the website's revenue. **Case 2:** An installment is paid ahead of time and there are ongoing obligations (i.e., $D_{it}^d > D_{it}^p$ and $\sum_{t=1}^{T_i} y_{it} < T_i$), or the last installment ($t = T_i$) is paid a few (within 30) days ahead of time (i.e., $\left(D_{i,T_i}^d - D_{i,T_i}^p\right) \leq 30$ and $\sum_{t=1}^{T_i} y_{it} = T_i$). In this case, the website not only can gain the interest revenue, but also can allocate the prepayment amount to the next loan. For simplicity, we assume that this extra benefit is linearly related to the time gap between the prepayment date and the due date with a weight γ assumed to be 1.15, which is the average ratio of gains to the repaid amount.[19] **Case 3:** The loan is paid off (more than 30 days) ahead of the loan due date (i.e., $\left(D_{i,T_i}^d - D_{i,T_i}^p\right) > 30$ and $\sum_{t=1}^{T_i} y_{it} = T_i$), As a reward incentive, a proportion (denoted as $r = 0.5$) of the interest from these prepaid installments would be waived. For example, if a 6-month loan is paid off in the fourth month, the loan interest for the fifth and sixth installments would be partially waived. **Case 4**: An installment is delinquent (i.e., $D_{it}^d > D_{it}^p$) In addition to the revenues from loan interest, the focal website charges a fine (with a daily rate as α) for late payment. In the present study, α equals 4.5‰. Mathematically, we have:

$$Revenue_{it} = \begin{cases} A_{it} \cdot IR_{it}, & D_{it}^d = D_{it}^p; \\ A_{it} \cdot IR_{it} \cdot \frac{\gamma\left(D_{it}^d - D_{it}^p\right)}{30}, & \left(D_{it}^d > D_{it}^p\ and\ \sum_{t=1}^{T_i} y_{it} < T_i\right)\ or\ \left(\left(D_{i,T_i}^d - D_{i,T_i}^p\right) \leq 30\ and\ \sum_{t=1}^{T_i} y_{it} = T_i\right); \\ A_{it} \cdot rIR_{it} \cdot \frac{\gamma\left(D_{it}^d - D_{it}^p\right)}{30}, & \left(D_{it}^d - D_{it}^p\right) > 30\ and\ \sum_{t=1}^{T_i} y_{it} = T_i\,; \\ A_{it}\left(IR_{it} + \alpha\left(D_{it}^p - D_{it}^d\right)\right), & D_{it}^d < D_{it}^p. \end{cases} \quad \text{(A1)}$$

Step 2: Cost Calculation

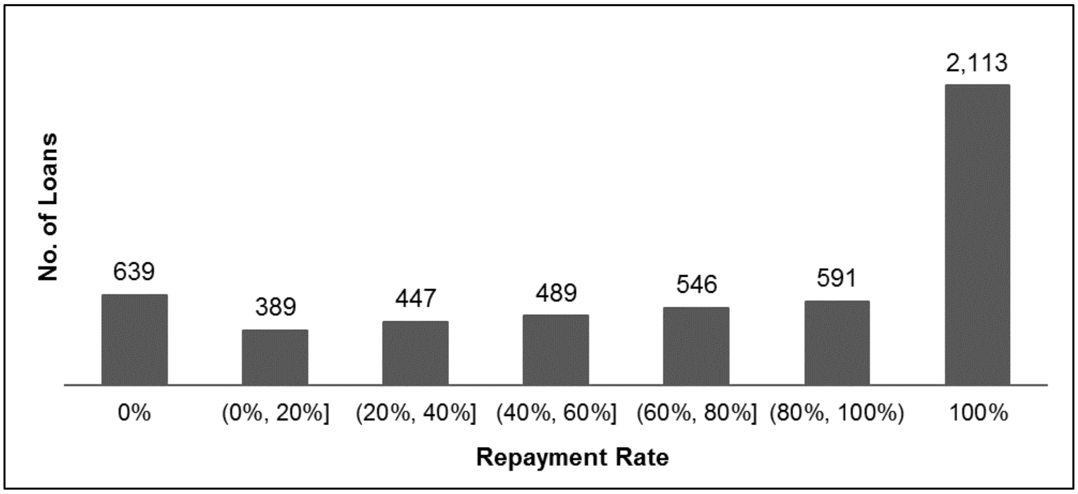
The cost is from the default case ($Cost_{it} = \delta \cdot A_{it}$). We define δ to capture the opportunity cost of unpaid installments ($\delta \geq 1$). We also set δ to 1.15 as per the average ratio of gains to the repaid amount. Because the fine charged for late payment includes a penalty, an operational cost for debt management (e.g., debt collection), and an opportunity cost for reuse in a new loan, we do not specially consider the opportunity cost from delinquency case.

Step 3: Loan Profit Calculation

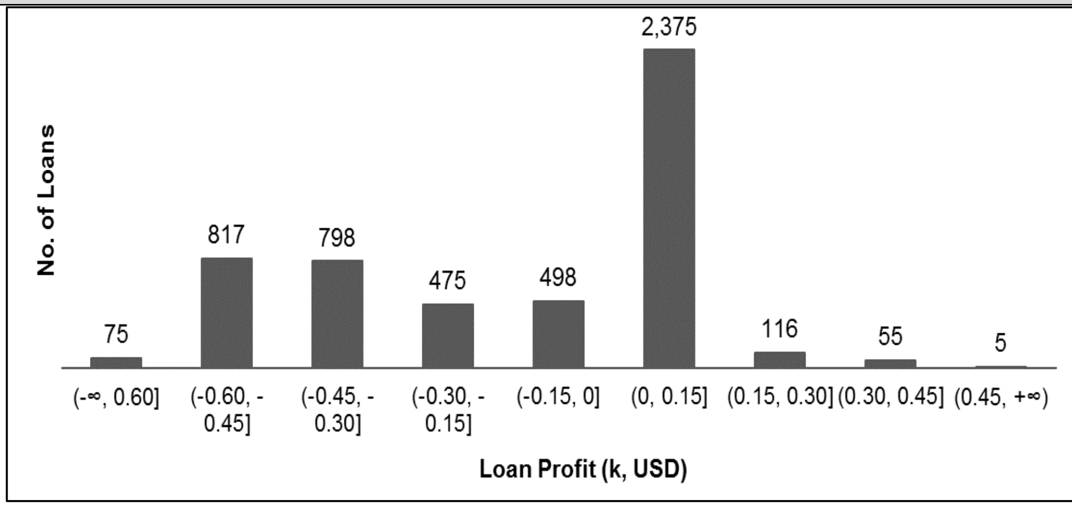
Based on the above revenue and cost definitions, we have the profit from loan i's installment t as: $Profit_i = y_{it} \cdot Revenue_{it} - (1 - y_{it}) \cdot Cost_{it}$.

The total profit of loan i is: $Profit_i = \sum_{t=1}^{T_i} Profit_{it}$.

[19] At the focal microloan website, the money paid back can always be lent out again soon.



(a) Repayment rate



(b) Loan profits (k, USD)

Figure B1. Distribution of Repayment Rates and Loan Profits

Appendix C

Prediction Details and Results

We implemented various widely accepted machine learning models, including logistic and linear regression (L&R), support vector machine (SVM), *k*-nearest neighbor (*k*-NN), multilevel perceptron (MLP), random forest (RF) and XGBoost. Notably, a self-selection issue might arise when applying social media features to train models directly with a microblogger subsample. Thus, following the two-stage framework proposed by Heckman (1979), we applied conventional features, online activity features, and mobile activity features to train the microblog usage decision ($if\_microblog$, 1 = *yes*, 0 = *no*) in the first stage and to obtain its inverse Mills ratio. Then, we included it in the second stage of credit risk prediction to compensate for self-selection bias. Because of the unbalanced distribution of each class for the categorical credit risk indicators, we implemented the oversampling strategy for minority classes in order to balance the trade-off among all classes. To avoid overfitting issues, we implemented feature selections via the L1-norm-based regularized sparse model before training any machine learning models.

We randomly partitioned our sample into two parts; approximately two thirds (3,476 loans) served as the training and validation sample, and the remaining one third (1,738 loans) served as the testing sample. We then applied 10-fold cross-validation to train the various models based on the training and validation sample and evaluated the prediction performance using the testing sample. Following Cui et al. (2018), we used grid search to choose the hyperparameter value yielding the best performance. We considered several commonly adopted metrics to evaluate prediction performance. Specifically, for the multiclass categorical risk indicator ($delinquent/default$), we considered precision, recall, and F1 score; for the numerical risk indicators ($repayment\ rate$ and $loan\ profit$), we considered mean absolute error (MAE), root mean squared error (RMSE), and *R*-squared (R^2). Table C1 reports the prediction performances of the proposed categorical credit risk indicators for the testing set. Overall, the results with different evaluation metrics and evaluation outcomes show consistency. As suggested by Abbasi et al. (2012), we conducted paired *t*-tests to compare the performances of the different alternative feature sets against the conventional feature set (Tables C2 and C3).

Table C1. Prediction Performances of Credit Risk Indicators (Delinquent/Default)

Model	***L&R***			***SVM***			***k-NN***		
Feature set	**Precision**	**Recall**	**F1 score**	**Precision**	**Recall**	**F1 score**	**Precision**	**Recall**	**F1 score**
Fc	0.357	0.362	0.359	0.360	0.363	0.361	0.338	0.319	0.328
Fo	0.401	0.407	0.404	0.397	0.402	0.399	0.390	0.371	0.380
Fm	0.536	0.583	0.559	0.548	0.587	0.567	0.555	0.550	0.552
Fs	0.525	0.569	0.546	0.508	0.569	0.537	0.541	0.490	0.514
Fc∪Fo∪Fm	0.538	0.586	0.561	0.556	0.590	0.572	0.559	0.554	0.556
Fc∪Fo∪Fm∪Fs	0.540	0.588	0.563	0.557	0.592	0.574	0.560	0.555	0.557
Model	***MLP***			***RF***			***XGBoost***		
Feature set	**Precision**	**Recall**	**F1 score**	**Precision**	**Recall**	**F1 score**	**Precision**	**Recall**	**F1 score**
Fc	0.351	0.350	0.350	0.424	0.361	0.390	**0.425**	**0.370**	**0.396**
Fo	0.408	0.420	0.414	0.449	0.418	0.433	**0.479**	**0.438**	**0.458**
Fm	0.725	0.607	0.661	0.627	0.613	0.620	**0.737**	**0.638**	**0.684**
Fs	**0.663**	0.583	**0.620**	0.564	0.560	0.562	0.549	**0.585**	0.566
Fc∪Fo∪Fm	0.728	0.616	0.667	0.625	0.619	0.622	**0.738**	**0.641**	**0.686**
Fc∪Fo∪Fm∪Fs	0.729	0.617	0.668	0.627	0.622	0.624	**0.740**	**0.643**	**0.688**

Note: Values in bold indicate the highest values for the various models.

Table C2. P-values of Pair-Wise *T*-tests for Alternative Feature Categories vs. Conventional Features (based on XGBoost)

Outcome	**Metric**	**Fo *vs.* Fc**	**Fm *vs.* Fc**	**Fs *vs.* Fc**	**Fc∪Fo∪Fm *vs.* Fc**	**Fc∪Fo∪Fm∪Fs *vs.* Fc**
$delinquent/default$	Precision	0.078*	< 0.001***	0.027**	< 0.001***	< 0.001***
	Recall	0.073*	0.003***	0.009***	0.003***	0.002***
	F1 score	0.075*	< 0.001***	0.009***	< 0.001***	< 0.001***
$repayment\ rate$	MAE	0.096*	< 0.001***	< 0.001***	< 0.001***	< 0.001***
	RMSE	0.082*	< 0.001***	< 0.001***	< 0.001***	< 0.001***
	R^2	0.007***	< 0.001***	< 0.001***	< 0.001***	< 0.001***
$loan\ profit$	MAE	0.105	< 0.001***	< 0.001***	< 0.001***	< 0.001***
	RMSE	0.108	< 0.001***	0.007***	< 0.001***	< 0.001***
	R^2	0.003***	< 0.001***	< 0.001***	< 0.001***	< 0.001***

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C3. P-values of Pair-Wise *T*-tests for Alternative Feature Categories vs. Conventional Features in Profit Analysis (based on XGBoost, Loan Approval Rate 45%)

Outcome metric for prediction	Fo *vs.* Fc	Fm *vs.* Fc	Fs *vs.* Fc	Fc∪Fo∪Fm *vs.* Fc	Fc∪Fo∪Fm∪Fs *vs.* Fc
$default$	0.053*	0.009***	0.022**	< 0.001***	< 0.001***
$delinquent\ but\ not\ in\ default$	0.088*	0.009***	0.030**	0.002***	< 0.001***
$Repayment\ rate$	0.097*	< 0.001***	0.028**	< 0.001***	< 0.001***
$Loan\ profit$	0.110	0.009***	0.028**	< 0.001***	< 0.001***

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We also ran the credit risk predictions using different sets of variables across the two types of loan purposes (Table C4). We found: (1) consistent result patterns with our main analyses based on the combined sample, and (2) overall similar prediction performances of most feature sets including the cellphone usage and mobility features (*Fm*) between the two loan purposes. This was due to the fact that these features reflect borrowers' relatively long-term behavioral tendencies rather than instant behaviors around the loan application date, and/or the fact that only a small proportion of borrowers in the sample had peculiar cellphone usage for emergencies. In addition, we found that the predictive power of online consumption features (*Fo*) became weaker for the consumption loan purpose group than for the other group (F1 scores 0.439 vs. 0.462). One plausible explanation is that since the borrowers in the consumption loan purpose group may tend to have overall more similar consumption preferences/activities compared with the entire borrower sample, *Fo* becomes less discriminative for screening creditworthiness.

Table C4. Prediction Performance of Delinquency/Default Indicators across Loan Purposes (Based on Microblogger Subsample and XGBoost)

Model	Loan purpose = 0 (for emergencies)			Loan purpose = 1 (for consumption)			Loan purpose (all)		
Feature set	***Precision***	***Recall***	***F1 score***	***Precision***	***Recall***	***F1 score***	***Precision***	***Recall***	***F1 score***
Fc	0.423	0.366	0.392	0.427	0.371	0.397	0.425	0.370	0.396
Fo	0.485	0.441	0.462	0.462	0.419	0.439	0.479	0.438	0.458
Fm	0.740	0.640	0.686	0.729	0.636	0.679	0.737	0.638	0.684
Fs	0.551	0.579	0.565	0.553	0.584	0.568	0.549	0.585	0.566
Fc∪Fo∪Fm	0.741	0.643	0.689	0.735	0.634	0.681	0.738	0.641	0.686
Fc∪Fo∪Fm∪Fs	0.741	0.644	0.689	0.735	0.634	0.681	0.740	0.643	0.688

Benefit-Cost Analysis

Finally, we performed a simple back-of-the-envelope calculation on the cost of replicating our study by other companies, as shown below (note that the calculation does not include the cost of acquiring alternative datasets because, for a company, this cost is incurred prior to their determination to conduct an experiment).

Cost

The experimental analysis involves approximately 100 features across four data sets, and after preliminary feature selection, the 22 most important features remain useful for prediction. Thus, for regular machine learning prediction processing (i.e., having at least a training set and an out-of-sample testing set division), the sample size should be no smaller than 220 (i.e., 10 times the number of features/parameters) (Hua et al. 2005) in order to obtain effective prediction outcomes and avoid overfitting. The regular and optimal approval rate on the website is 45% and the profit per loan is 11 USD (Page 24). In contrast, our experimental design approves all of the loan applications without selection, and the average payoff of the 55% of loans that should have been excluded if the experiment were not conducted is approximately -195 USD. Therefore, the lowest cost of conducting the experiment is -23,600 USD (= -195 × 220 × 55%). Note that the website's other costs such as labor and managerial costs are compensated by the service fee, which is absorbed in the calculation of loan profit.

Benefit

The benefit mostly comes from the improved profitability after conducting the experiment and figuring out the most effective feature sets. Therefore, taking the common practice of applying only conventional features (*Fc*) and an approved sample for model training as the benchmark, we compare it with the case of leveraging the optimal feature sets. According to our analysis in Figure 1, the most useful alternative feature set (i.e., *Fm*) generates incremental gains of approximately 4,440USD (= 14,900 - 10,460) for the 243 approved loans, i.e., 18 USD per loan (compared with the original loan payoff of 11 USD, the incremental gains by the experiment are rather intriguing). Therefore, considering no opportunity cost, the website needs to issue approximately 1,300 (= 23,600/18) loans after the experiment to offset the cost of conducting the experiment.

Appendix D

Approved Sample-Based Prediction and Financial Equality

Table D1. Comparison of Conventional Features

Feature	Mean		*SD*	
	Approved sample	***Full sample***	***Approved sample***	***Full sample***
Gender	0.28	0.20	0.46	0.41
Age	25.10	24.33	3.87	3.40
City DPI	6,331	5,102	2,634	2,111
Monthly income level	4.65	3.22	1.55	1.41
Loan-to-income ratio	1.10	1.60	1.51	1.42
Homeownership	0.22	0.18	0.40	0.40
Marital status	0.12	0.10	0.32	0.33
# Children	0.08	0.10	0.31	0.37
Education level	4.16	3.33	0.73	0.64
# Registered contacts	3.16	2.46	0.86	0.85
Loan amount	477	466	87	83
Loan term	5.72	5.68	1.82	1.79
Loan interest rate	13.69	13.70	0.63	0.63
Loan purpose	0.50	0.50	0.49	0.50
If_other loan (history)	0.31	0.32	0.56	0.46
# Contacted microloan websites	1.10	1.38	1.77	1.92
# Contacts with microloan websites	2.86	4.08	7.20	8.57
# Call out to microloan websites	2.70	2.68	6.75	5.98
# Call in from microloan websites	0.11	1.48	1.11	2.36
# Registered microloan companies	5.83	8.25	1.80	2.08
If_other loan (current)	0.01	0.002	0.12	0.04
# Loans borrowed on microloan websites	0.29	0.09	0.47	0.47
If_delinquent on other microloan websites	0.07	0.35	0.21	0.46
# Defaults_first-order contacts	0.06	0.07	0.36	0.40
If_credit card	0.10	0.05	0.29	0.24
If_pay credit card regularly	0.06	0.03	0.24	0.17
Type of occupation	2.48	1.98	1.40	1.21
If_insurance	0.64	0.56	0.46	0.51

Note: The full sample in this table is a randomly drawn subsample from the original training sample of the experimental data. The sample size of both the approved sample and the full sample is 1,043.

Appendix E

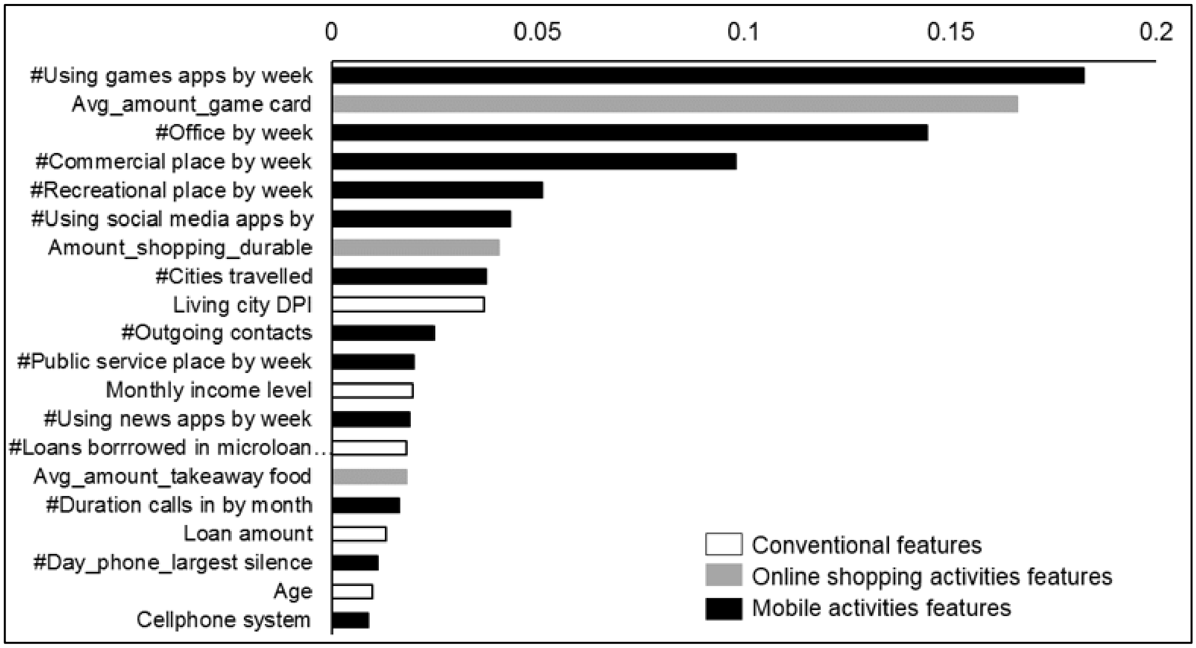
Explanation Detection

We explored, at a more granular level, which features are the most relevant to financial credit risk prediction. For this purpose, we implemented a permutation feature attribution method (Fisher et al., 2018). The permutation is based on the resulting drop in the accuracy of the model when a single feature is randomly permutated in the test data set. We ranked the normalized permutation importance of all of the features. Figure E1 displays the top 20 most important features for the whole sample. As is consistent with previous findings, Figure E1(a) shows the importance of the features in the mobile activities category as well as in the online activities category. Interestingly, game-relevant behaviors (i.e., frequency of using game apps and average purchase amount of game card) played significant roles in predicting borrowers' credit risk. Game-relevant behaviors reflect self-indulgence to a large extent (Kim et al., 2008), and users who frequently consume games may be more likely to engage in excessive consumption and to fail to maintain a sound financial plan. In addition to these features, borrowers' mobility trajectories, including appearance frequencies in official buildings, commercial places, and recreational places, were also important. Frequently and/or regularly appearing in official areas indicates that users are maintaining steady work and income sources, while frequently shopping in commercial and recreational places may result in daily financial constraints (Mehrotra et al., 2017). These features thus are highly associated with individuals' repayment performance for microloans. Yet, in general, the conventional features (e.g., city DPI and monthly income level) are less important than the new alternative sources of information. This implies that the seemingly straightforward features measuring users' economic capacity are less relevant to their credit behaviors than certain alternative features. One possible explanation is that common experience indicates that users' subjective sense of economic pressure might not always be consistent with their objective economic capacity. Psychosocial characteristics (e.g., economic strain and personality) as reflected by extended alternative information sometimes outweigh economic capacities reflected by income and loan histories with respect to individuals' decision to fulfill financial obligations as well as their credit behaviors (Lu et al., 2020). When considering social media features for the social media activities samples, we learned from Figure E1(b) that several social media activity characteristics (i.e., the number of fans a borrower maintains, the sentiment valences of the originally generated messages, and the number of "likes" received in her microblog) are likewise quite important for prediction of financial credit risk. Social media presence and posts on social media also show personalities and psychosocial status, including socialization strains, which have been corroborated as being associated with individuals' financial behaviors (Ge et al., 2017; Lu et al., 2020).

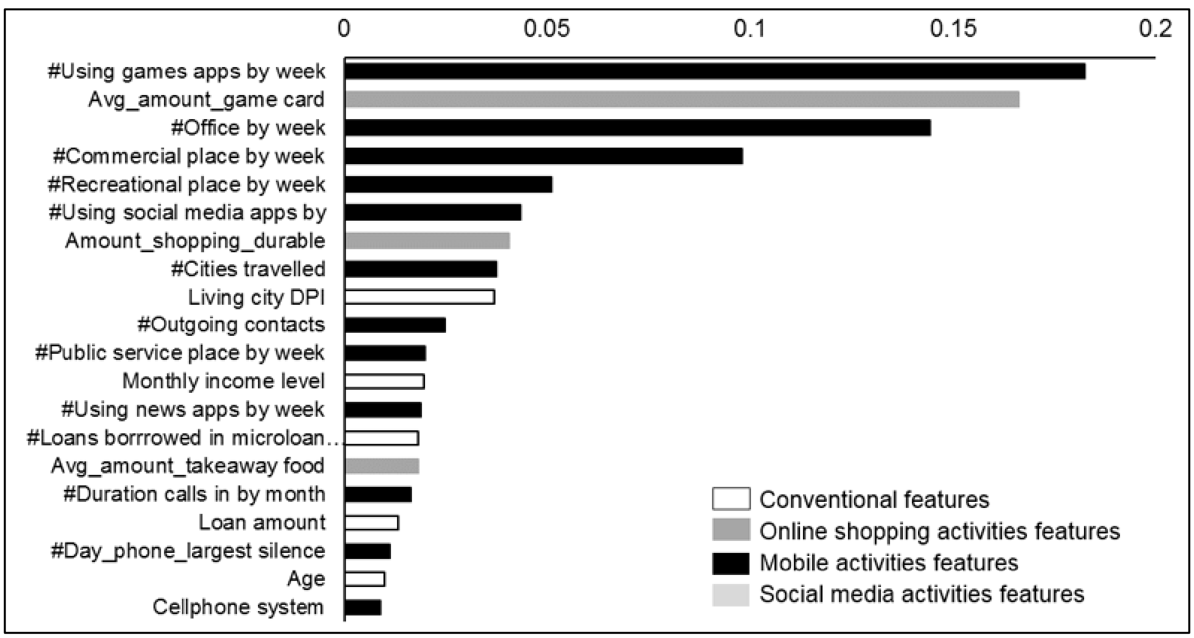
Table E1. Correlations between Alternative and Sensitive User Features

Alternative features		Sensitive User Features (Fc)				
		City DPI	*Monthly income level*	*Loan-to-income ratio*	*Education level*	*Home-ownership*
Fo	Avg_amount_game card	-0.004	0.001	-0.003	-0.041	-0.007
	Diversity_shopping_durable	0.078	0.149	-0.017	0.047	0.013
	Amount_shopping_durable	0.357	0.401	-0.438	0.074	0.062
	# Order_shopping_durable	0.296	0.413	-0.316	0.122	0.223
	Amount_transfer out	0.045	0.188	-0.236	0.051	0.016
	Amount_shopping_virtual	0.160	0.262	-0.072	0.359	0.034
Fm	# Office by week	0.019	0.002	0.006	0.039	0.010
	# Commercial place by week	0.017	0.026	-0.011	0.051	0.000
	# Recreational place by week	0.002	0.020	-0.011	-0.021	0.005
	# Using games apps by week	0.008	-0.012	-0.007	-0.033	-0.020
	# Outgoing contacts	0.000	0.066	-0.012	0.040	0.008
	# Calls in by month	-0.072	0.016	0.008	-0.061	0.046
Fs	# Fans in microblog	-0.008	0.011	-0.002	-0.004	0.014
	Sentiment valence of generated messages	-0.012	-0.002	0.003	-0.016	-0.009
	# Likes in microblog	0.010	0.004	-0.007	0.040	0.016
	# Comments in microblog	0.013	0.013	-0.017	0.043	0.025
	Sentiment valance	-0.011	0.003	0.000	-0.014	-0.007
	# Generated messages in microblog	0.016	0.008	-0.008	0.038	0.021

Note: These alternative features are those ranking high in feature-importance analyses. Shadow cells indicate that the correlation value is larger than 0.2 and is significantly different from 0.



(a) Full Sample (without Social Media Features)



(b) Microblogger Subsample (with Social Media Features)

Figure E1. Feature Importance of Categorical Risk Indicators (Top 20)

Appendix F

Geographical Bias Issue

Certain geographic areas might be more likely to produce certain types of alternative data, which might result in the model being biased toward/against some of these areas. To test if this conjecture holds, we performed two analyses. First, we partitioned our sample users into five subsets according to the cities in which they live: northern, southern, eastern, western, and middle areas of the focal country. It is well recognized that people across the five areas of the focal country generally present somewhat distinct daily life habits and cultural thoughts due to the different geographic features, weather and climates, and cultural traditions. The sample sizes for the five subsets were 1,410, 822, 954, 1,113, and 915. Then, we fixed the out-of-sample testing set containing the samples from all five areas. In model training, rather than applying a purely random fivefold cross-validation, we used the subsample of one geographic area as the validation set, the subsample of the other four areas as the training set, and obtained the prediction outcomes on the out-of-sample testing set. We permutated the validation set with the five subsamples of the different areas and averaged the prediction performance. We applied consistent operations with (1) XGBoost prediction on Default, (2) different feature sets, and (3) full sample *vs.* approved sample, in our main analyses. We believe that this operationalization helped us to examine whether our data sets and model training would be heavily biased toward/against samples of some areas. The prediction results in Tables F1 and F2 are in line with our main analyses, suggesting that, in our context, the concern about geographic clustering of data is trivial.

Table F1. Prediction Performance Comparison (Based on Dictions with Predefined Geographic Areas and XGBoost)

Feature set	Delinquent/default (F1 score)		
	Approved- sample	***Full- sample***	***Bias (%)***
Fc	0.160	0.371	56.87
Fo	0.339	0.428	20.79
Fm	0.524	0.650	19.38
Fs	0.340	0.518	34.36
Fc∪Fo∪Fm	0.529	0.660	19.85
Fc∪Fo∪Fm∪Fs	0.529	0.659	19.88

Note: Bias = (Full sample-based - Approved sample-based) / Full sample-based.

Table F2. Comparisons of Top 45% Best-Approved Loans (Based on Predictions with Predefined Geographic Areas, Microblogger Subsample, XGBoost, Default-Based Prediction)

		Loan selection strategy			
		Group A	***Group B***	***Group C***	***Group D***
		Approved sample with Fc	**Full sample with Fc**	**Approved sample with Fc∪Fo∪Fm∪Fs**	**Full sample with Fc∪Fo∪Fm∪Fs**
# (ratio) of overlap loans to Group D		148 (60.91%)	171 (70.37%)	204 (83.95%)	243 (100%)
Means of samples of conventional features of unique borrowers	City DPI	7,810.41	7,331.20	6,775.15	6,502.10
	Monthly income level	5.51	5.28	4.56	4.36
	Loan-to-income ratio	1.16	1.32	1.42	1.42
	Education level	4.26	4.20	3.96	3.85
	Homeownership	0.58	0.52	0.44	0.36

Note: Sample features are those showing significantly different mean values across groups.

To further tease out the issue, we next followed Meyer and Pebesma (2021) (using their developed code package *CAST*) to perform a (spatial) prediction model incorporating the "area of applicability" (AOA); this is the area wherein the model is enabled to learn about relationships based on the training data and wherein the estimated cross-validation performance holds. This model first estimates the "dissimilarity index" (DI) that is based on the minimum distance to the training data in the multidimensional predictor space. Predictors are weighted based on the variable importance of the machine learning algorithm used for model training. The model then derives the AOA by considering the distance from the new data in the predictor variable space to the data points used for model training. Specifically, the AOA is derived by applying a DI threshold, which is the (outlier-removed) maximum DI of the cross-validated training data (Meyer et al., 2020). As such, we also applied, to XGBoost, the AOA computation across the sample users (i.e., data points) of the different geographic areas (provinces). After this amendment, we replicated our predictions on the default and presented predictions for the AOA only. The results in Tables F3 and F4 are also similar to our main analyses and help further mitigate the concern about the geographic clustering of data.

Table F3. Prediction Performance Comparison (based on Predictions for AOA and XGBoost)

Feature set	Delinquent/default (F1 score)		
	Approved sample	***Full sample***	***Bias (%)***
Fc	0.180	0.409	55.99
Fo	0.359	0.466	22.96
Fm	0.552	0.690	20.00
Fs	0.388	0.581	33.22
Fc∪Fo∪Fm	0.558	0.699	20.17
Fc∪Fo∪Fm∪Fs	0.558	0.700	20.29

Note: Bias = (Full sample-based - Approved sample-based)/Full sample-based. The AOA thresholds for the predictions range from 0.44 to 0.67.

Table F4. Comparisons of Top 45% Best-Approved Loans (based on Predictions for AOA, Microblogger Subsample, XGBoost, Default-based Prediction)

		Loan selection strategy			
		Group A	***Group B***	***Group C***	***Group D***
		Approved sample with *Fc*	**Full sample with *Fc***	**Approved sample with *Fc*∪*Fo*∪*Fm*∪*Fs***	**Full sample with *Fc*∪*Fo*∪*Fm*∪*Fs***
# (ratio) of overlap loans to Group D		160 (65.84%)	180 (74.07%)	216 (88.89%)	243 (100%)
Means of samples of conventional features of unique borrowers	City DPI	7,780.20	7,587.19	6,432.12	6,191.54
	Monthly income level	5.40	5.21	4.37	4.33
	Loan-to-income ratio	1.11	1.21	1.36	1.39
	Education level	4.28	4.15	4.00	3.94
	Homeownership	0.54	0.48	0.41	0.36

Note: Sample features are those showing significantly different mean values across groups.

Appendix G

Measurement of Equalized Opportunity

To formally test whether the full sample and certain types of alternative feature sets can alleviate bias and promote fairness, we examine the fairness of the algorithm with respect to protected demographic indicators. Specifically, we apply the fairness criterion of "equalized opportunity," namely that positive outcomes should be independent of the protected attribute (Teodorescu et al., 2021). Let A be the sensitive demographic feature ($A = 1, 2, \dots, n$), and $Y = 1$ and $\hat{Y} = 1$ be the correct and actual positive outcomes (i.e., in our context, a loan application being approved), respectively, namely the cases of being correctly approved (non-default case) and being actually approved, respectively; then, equalized opportunity means $p(\hat{Y} = 1 | A = 1, Y = 1) = p(\hat{Y} = 1 | A = 2, Y = 1) = \cdots = p(\hat{Y} = 1 | A = n, Y = 1)$. In this vein, taking the education level and monthly income level as examples, we examined the loan approval ratios of borrowers of different education levels and monthly income levels in our testing set under different scenarios. Tables G1 and G2 indicate that compared with the prediction based on the approved sample and conventional features only, the probabilities of non-defaulted borrowers of lower education levels and lower monthly income levels being approved become more similar to those of higher education levels and higher monthly income levels. For example, for the actual non-defaulted borrowers who only have a middle school education or less, only 36% can be approved by the platform if only approved sample and *Fc* are applied; in contrast, 71% of the actual non-defaulted postgraduate borrowers can be approved. Fortunately, when using the full sample and all feature sets ($Fc \cup Fo \cup Fm \cup Fs$), the approval ratio for the actual non-defaulted borrowers who have a middle school education or less reaches 72%. Recall that we have demonstrated that applying the full sample and alternative features yields higher default prediction accuracies and profits. Our findings suggest that applying either the full sample or alternative features can improve borrower selection accuracy while also improving financial fairness. Applying the full sample and alternative features simultaneously generates optimal results with respect to achieving both financial profitability and fairness. We also examined other sensitive features such as homeownership and city DPI and obtained consistent findings (these results are available upon request).

Table G1. Financial Equality Analysis of Borrowers of Different Education Levels

Education level, A = *n*	%Actually be approved for non-defaulters, $p(\hat{Y} = 1 \mid A = n, Y = 1)$, prediction with approved sample	%Actually be approved for non-defaulters, $p(\hat{Y} = 1 \mid A = n, Y = 1)$, prediction with full sample
(a) Prediction with Fc		
A = 1, middle school or less	0.36	0.49
A = 2, vocational school	0.45	0.59
A = 3, high school	0.49	0.66
A = 4, technical school	0.61	0.67
A = 5, undergraduate	0.66	0.72
A = 6, postgraduate	0.71	0.71
(b) Prediction with $Fc \cup Fo \cup Fm \cup Fs$		
Education level, A = *n*	**%Actually be approved for non-defaulters, $p(\hat{Y} = 1 \mid A = n, Y = 1)$, prediction with approved sample**	**%Actually be approved for non-defaulters, $p(\hat{Y} = 1 \mid A = n, Y = 1)$, prediction with full sample**
A = 1, middle school or less	0.54	0.72
A = 2, vocational school	0.62	0.79
A = 3, high school	0.66	0.80
A = 4, technical school	0.72	0.85
A = 5, undergraduate	0.81	0.90
A = 6, postgraduate	0.86	1.00

Note: A indicates education level, Y = 1 indicates the case of being correctly approved (i.e., non-default case), and $\hat{Y} = 1$ indicates the case of being actually approved.

Table G2. Financial Equality Analysis of Borrowers of Different Monthly Income Levels

(a) Prediction with Fc		
Monthly income level, A = *n*	**%Actually be approved for non-defaulters, p($\hat{Y}$ = 1|A = *n*, Y = 1), prediction with approved sample**	**%Actually be approved for non-defaulters, p($\hat{Y}$ = 1|A = *n*, Y = 1), prediction with full sample**
A = 1, 150 USD or less	0.46	0.56
A = 2, 150-300 USD	0.55	0.58
A = 3, 300-450 USD	0.64	0.70
A = 4, 450-600 USD	0.66	0.74
A = 5, 600-750 USD	0.76	0.73
A = 6, 750-900 USD	0.79	0.84
A = 7, 900 USD or more	0.81	0.85
(b) Prediction with Fc∪Fo∪Fm∪Fs		
Monthly income level, A = *n*	**%Actually be approved for non-defaulters, p($\hat{Y}$ = 1|A = *n*, Y = 1), prediction with approved sample**	**%Actually be approved for non-defaulters, p($\hat{Y}$ = 1|A = *n*, Y = 1), prediction with full sample**
A = 1, 150 USD or less	0.70	0.78
A = 2, 150-300 USD	0.75	0.84
A = 3, 300-450 USD	0.80	0.89
A = 4, 450-600 USD	0.82	0.94
A = 5, 600-750 USD	0.82	0.94
A = 6, 750-900 USD	0.93	1.00
A = 7, 900 USD or more	0.94	1.00

Note: A indicates education level, Y = 1 indicates the case of being correctly approved (i.e., non-default case), and $\hat{Y}$ = 1 indicates the case of being actually approved.

Appendix H

Canonical Correlation Analysis

Following Thompson (2000), we performed a canonical correlation analysis (CCA) to test the correlations between online shopping activity features and traditional demographic features as well as between mobile activity features and traditional demographic features. In CCA, the model projects the features of each feature set onto linear combinations of features, which are termed "projection features." The purpose of CCA is to maximize the correlations between projection features from two feature sets. The maximum correlation between two sets of projection features represents the information overlap between them, which is used to indicate the correlation between the two feature sets. As shown in Figure H1, the correlation between mobile activity features and traditional demographic features was only 0.07, whereas the correlation between online shopping activity features and traditional demographic features was as high as 0.22, which suggests the orthogonal relationship between mobile activity features and traditional demographic features.

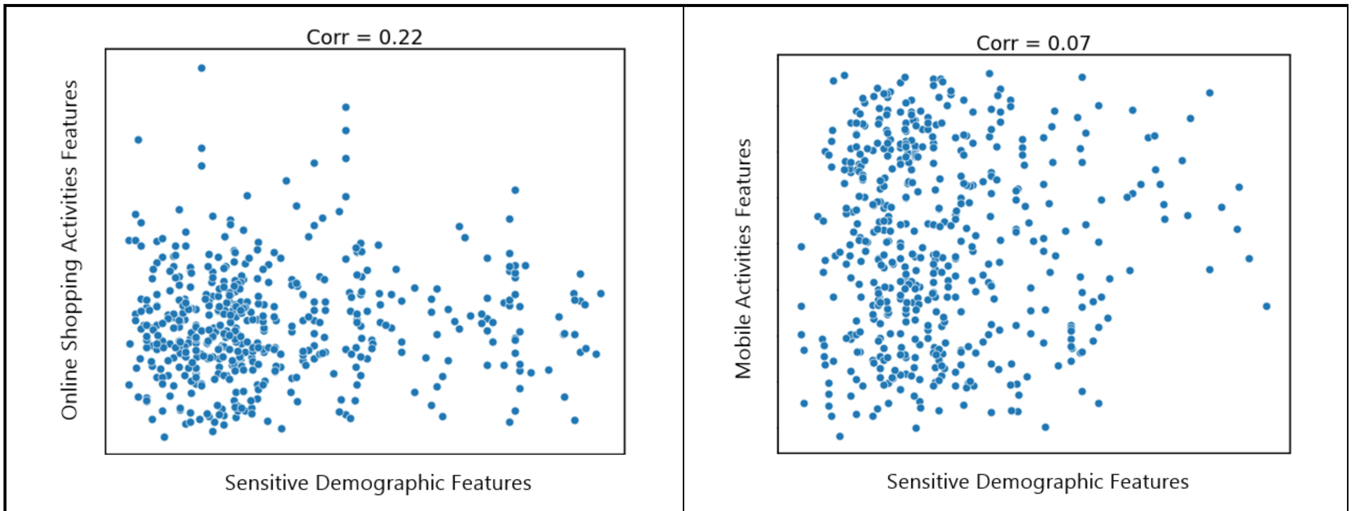


Figure H1. Scatter Plot of Values of Projection Features for Alternative Features and Traditional Demographic Features (Output by CCA)

Appendix I

Data Privacy Issue

Metalevel features are those extracted to describe the distribution of original features corresponding to each sample record, which is to say, the features extracted at the feature level. The principal tenet of metafeature-based analysis is to reconstruct borrowers' original feature space in a desensitized way but at the same time to maximize the retention of discriminative factors embedded in informative features. Taking the binary outcome variable $default$ as an example, default and non-default are the two instances of this variable. Essentially, the distribution of features fed to the prediction model varies over different instances. This is the main reason why default instances can be effectively identified from non-default instances (similarly to other outcome variables); and this endows the distribution of features with strong abilities to distinguish or classify among instances (c.f. Rauber et al., 2014). By constructing effective metalevel features, previous studies have achieved satisfactory performance of metafeature-based prediction in various applications such as fault diagnosis (Ciabattoni et al., 2017), audio signal processing (Sharma et al., 2020), and spatial filtering (Zhang et al., 2018). In particular, the distribution of feature values subject to each instance (i.e., in our case, each piece of loan record) can be regarded as sequence data (Wang et al. 2019). Therefore, the key is to extract appropriate metalevel features to efficiently describe the internal characteristics and fluctuation trends of sequence data.

We conducted a metafeature-based analysis across different alternative data categories to determine if our main findings still held. Specifically, based on the original features of each loan record for the experimental sample and the approved sample, we followed the canonical literature in order to extract metalevel features, as shown in Table I1. These features have been commonly used in prior studies (e.g., Rauber et al., 2014). We conducted identical operations as in the main financial profitability and equality analyses. We also applied $delinquent/default$, *repayment rate*, and *loan profit* as outcome variables. Regarding features, we first included metalevel features corresponding to the individual categories of conventional characteristics (*MFc*), online activities (*MFo*), mobile activities (*MFm*), and social media activities (*MFs*), respectively, among which *MFc* are the benchmark. Then, we combined features from the different categories ($MFc \cup MFo \cup MFm$ and $MFc \cup MFo \cup MFm \cup MFs$). Finally, we applied identical feature selection, parameter tuning, model training, and result comparison techniques.

Figure I1 indicates similar profit results by metafeature-based analysis. That is, the three kinds of alternative feature categories and the feature combinations would generate significantly larger profits for the website. Alternative features also help offset the website's economic losses caused by training sample bias. In fact, model training with alternative features based on an approved sample (11,710 USD) even yields higher profits than using simply conventional features with a full sample (9,850 USD). Furthermore, it is easy to learn from Table I2 that with metafeature-based analysis, alternative features are likewise able to include more low-risk borrowers traditionally considered to be "bad" users from less-developed areas who also have lower income and education levels.

Table I1. Extracted Metalevel Features

No.	Metalevel feature	Description	No.	Metalevel feature	Description
1-3	Quartiles (*First_quartile*, *Second_quartile*, *Third_quartile*)	–	10	Square root of amplitude (*SRA*)	$X_{SRA} = \left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lvert x_i\rvert}\right)^2$
4	Minimum value (*MIN*)	$X_{MIN} = Min(x_i)$	11	Skewness value (*SV*)	$X_{SV} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{\sigma}\right)^3$
5	Maximum value (*MAX*)	$X_{MAX} = Max(x_i)$	12	Kurtosis value (*KV*)	$X_{KV} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{\sigma}\right)^4$
6	Mean (*MEAN*)	$X_{MEAN} = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$	13	Crest factor (*CF*)	$X_{CF} = \frac{\max(\lvert x_i\rvert)}{\left(\frac{1}{N}\sum_{i=1}^{N} {x_i}^2\right)^{1/2}}$

7	Standard deviation (*STD*)	$X_{STD} = \sigma = \left(\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2\right)^{1/2}$	14	Impulse factor (*IF*)	$X_{CF} = \frac{\max(\lvert x_i \rvert)}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i \rvert}$
8	Peak-to-peak value (*PPV*)	$X_{PPV} = Max(x_i) - Min(x_i)$	15	Margin factor (*MF*)	$X_{CF} = \frac{\max(\lvert x_i \rvert)}{\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lvert x_i \rvert}\right)^2}$
9	Root mean square (*RMS*)	$X_{RMS} = \left(\frac{1}{N}\sum_{i=1}^{N}{x_i}^2\right)^{1/2}$	16	Shape factor (*SF*)	$X_{CF} = \frac{\left(\frac{1}{N}\sum_{i=1}^{N}{x_i}^2\right)^{1/2}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i \rvert}$

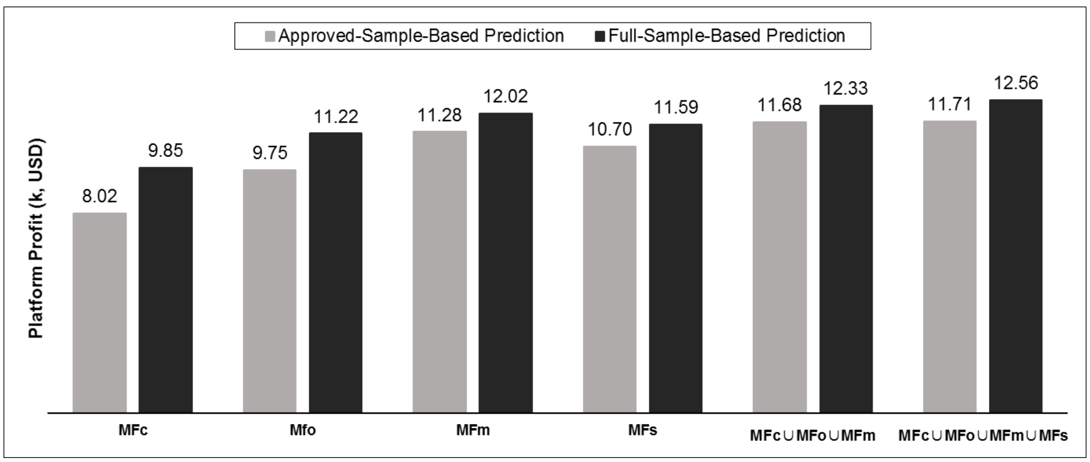


Figure I1. Profit Comparison (Based on Microblogger Subsample, XGBoost, Metafeature-Based Prediction on Default, Loan Approval Rate 45%)

Table I2. Comparisons of Top 45% Best-Approved Loans (based on Microblogger Subsample, XGBoost, Metafeature-based Prediction on Default)

		Loan approval strategy			
		Group 6	***Group 1***	***Group 10***	***Group 5***
		Approved sample-based prediction with *Fc*	**Full sample-based prediction with *Fc***	**Approved sample-based prediction with all features**	**Full sample-based prediction with all features**
# (ratio) of overlap loans to Group 5		129 (53.75%)	152 (63.33%)	194 (80.83%)	240 (100%)
Means of samples of conventional features of unique borrowers	City DPI	8,086.43	7,821.21	7,498.01	6,981.55
	Monthly income level	5.56	5.45	4.82	4.33
	Loan-to-income ratio	1.10	1.16	1.37	1.44
	Education level	4.40	4.29	4.13	3.82
	Homeownership	0.58	0.52	0.45	0.41

Note: Sample features are those showing significantly different mean values across groups.