

Johnen, Constantin; Mußhoff, Oliver

Article — Published Version

Digital credit and the gender gap in financial inclusion: Empirical evidence from Kenya

Journal of International Development

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Johnen, Constantin; Mußhoff, Oliver (2022) : Digital credit and the gender gap in financial inclusion: Empirical evidence from Kenya, Journal of International Development, ISSN 1099-1328, Wiley, Hoboken, NJ, Vol. 35, Iss. 2, pp. 272-295, <https://doi.org/10.1002/jid.3687>

This Version is available at:

<https://hdl.handle.net/10419/287809>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Digital credit and the gender gap in financial inclusion: Empirical evidence from Kenya

Constantin Johnen  | Oliver Mußhoff

Department of Agricultural Economics and Rural Development, University of Goettingen, Goettingen, Germany

Correspondence

Constantin Johnen, Department of Agricultural Economics and Rural Development, University of Goettingen, Goettingen, Germany.

Email: constantin.johnen@uni-goettingen.de

Abstract

Formal digital credit raises hopes to decrease the gender gap in financial inclusion. However, up until now, it remains unknown whether these hopes are justified. Using nationally representative household surveys from Kenya, the present study aims to fill this gap. We find strong indication that formal digital credit, contrasting to expectations, has led to an increase in the gender gap in financial inclusion. We further find indication that the pervasive gender gap in the formal digital credit market is largely attributable to gender differences in socio-economic variables in combination with a lack of contract term heterogeneity in that market. The paper suggests that policies to strengthen women's position in society and/or to encourage contract term heterogeneity in the formal digital credit market could decrease the gender gap in financial inclusion.

KEYWORDS

digital financial services, formal credit, gender equality, regulation

1 | INTRODUCTION

Increasing financial inclusion, typically defined as access to or usage of formal financial products, is largely believed to be an important driver of economic growth (Demirgüç-Kunt & Klapper, 2013; Hannig & Jansen, 2010). A main focus within the stream of the financial inclusion literature is understanding the pervasive gender gap in financial inclusion through formal credit markets (Adegbite & Machethe, 2020; Ghosh & Vinod, 2017; Guerin, 2006; Swamy, 2014), causing thriving discussions on how to increase financial inclusion in general and for women in

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Journal of International Development* published by John Wiley & Sons Ltd.

particular (Banerjee et al., 2015; Bharadwaj & Suri, 2020; Ghosh & Vinod, 2017). A new innovation, formal digital credit, raises hopes to increase financial inclusion, especially for women, by leveraging the widely accessible mobile money network, which already counts 1.2 billion subscriptions worldwide (Andersson-Manjang & Naghavi, 2021).

Whether these hopes are justified remains largely unknown. In a recent study, Bharadwaj and Suri (2020) conclude that digital banking (which includes digital credit) holds its promise as an important source of financial inclusion. This finding is corroborated by recent studies which show that digital credit increases the total share of (formal) borrowers (Bharadwaj & Suri, 2020; Brailovskaya et al., 2021; Johnen et al., 2021). However, whether men and women profit equally from increases in financial inclusion through formal digital credit usage is not yet determined. In fact, early reports about digital credit in Kenya show that borrowers are more likely to be male (Gubbins & Totolo, 2018). However, these reports do not distinguish between formal digital credit: that is, at least one lending party is a regulated bank and semi-formal digital credit; that is, no lending party is a regulated bank. This distinction is relevant as only users of the former are considered financially included (Demirgüç-Kunt & Klapper, 2013). In addition, it appears conceivable that average formal digital credit user characteristics could change over time and might hence have changed from the reports about early digital credit users. In fact, literature about other digital financial services shows that initial gender differences in the usage of these services decrease over time (Jack & Suri, 2011).

In contrast, recent studies indicate that such a development might not take place in the formal digital credit market. Formal digital credit is inherently different from other digital financial services, as users need to qualify for usage. There is a growing body of literature which shows that algorithms assign systematically lower digital credit scores to women (Brailovskaya et al., 2021). These gender differences in credit scores might result from socio-economic differences, which for instance might cause different phone using patterns (Björkegren & Grissen, 2018). The mere existence of differing credit scores in itself does not necessarily imply a gender gap in formal digital credit users. Low credit scores can, in theory, be compensated through adjustments in contract terms, such as higher interest rates (Hoff & Stiglitz, 1990). However, in many low-and middle income countries, available formal digital credit contract terms are rather homogenous and inflexible. The reason for that is that the formal digital credit market is often dominated by very few mobile network operator-bank partnerships who offer a formal digital credit product with the same interest rate and repayment period for everyone (Blechman, 2016; Francis et al., 2017). The contract terms lenders are able to offer can additionally be delimited by market regulations, such as interest rate caps. This might lead to a further restriction in available contract terms. In Kenya for instance a single loan called M-Shwari dominates the formal digital credit market (Suri et al., 2021). Every M-Shwari loan has a fixed interest rate of 7.5% per month and a fixed repayment period of 1 month; contract terms are merely heterogeneous in credit size (Suri et al., 2021). This lack of heterogeneity in available contract terms might cause pervasive credit rationing, if none of the few available contract terms can be offered below a certain credit score (Hoff & Stiglitz, 1990). Systematically lower credit scores of women could therefore translate into a pervasive gender gap in the formal digital credit market.

It remains unknown whether initial gender differences in formal digital credit usage are indeed pervasive or, similar to other digital financial services, decrease over time. Consequently, there is also still very limited knowledge about why gender differences might persist over time. The present study addresses these issues. Therefore, our two central research questions are as follows: First, does formal digital credit decrease the gender gap in financial inclusion over time? Financial inclusion is defined as total formal credit usage, consisting of both formal digital and conventional credit. It is important to analyse both forms of formal credit usage to detect potential substitution effects. Second, what explains gender differences in formal digital credit usage? To answer the research questions, we use the nationally representative Kenya FinAccess household surveys from 2009, 2012, 2015 and 2018, which entail a wide range of socio-economic characteristics and formal digital credit usage. More precisely, to answer the first research question, we analyse the changes in the gender gap concerning financial inclusion and distinguish the relative contribution of formal digital and formal conventional credit. To answer the second research question, we conduct logistic regressions based on the latest survey round, in order to understand whether socio-economic gender differences explain the gender gap in formal digital credit usage. We compare the results to the association between socio-economic characteristics and credit usage in the semi-formal digital credit market, where a substantially larger number of lenders compete for

borrowers with different credit products. This comparison yields some further insight whether differences in market specific characteristics, such as contract term heterogeneity, are associated with the relationship between socio-economic variables and digital credit usage. Kenya is a particularly telling country for this investigation as formal digital credit was already introduced in 2012, which enables the authors to observe changes over time.

Our results show considerable, statistically significant and persistently lower user shares of women in the formal digital credit market. The differences in the formal digital credit market have further translated into an overall statistically significantly larger gender gap in the total formal credit market since the introduction of formal digital credit, from 4.1 percentage points in 2012 (just before digital credit was offered) to 7.8 percentage points in 2018. We find that gender differences in socio-economic variables such as income and education explain a large share of the gender gap in the formal digital credit market. In contrast, we do not find any gender differences in the unregulated semi-formal digital credit market (where the available contract terms are much more heterogeneous). We interpret this finding, in accordance with economic literature, as suggestive evidence that socio-economic differences explain the gender gap in the formal digital credit market partially due to the lack of suitable contract terms for women in that market.

By answering our research questions, this study contributes to several strands of literature. First, this study largely contributes to literature investigating the pervasive gender gap in financial inclusion. This stream of literature identifies several obstacles for women in the formal conventional credit market, foremost long-travel times to banks and collateral requirements (Baydas et al., 1994; Fletschner & Kenney, 2014). In the present study, we are the first to investigate the gender gap in formal digital credit, a type of credit that renders the aforementioned obstacles obsolete. We show that despite these promising characteristics, formal digital credit has largely contributed to a statistically significant increase in the gender gap in financial inclusion.

Second, we further contribute to the extensive body of literature that investigates adoption and usage of digital financial services. This stream of literature finds that initial gender differences in the usage of digital financial services decrease over time (Demombynes & Thegeya, 2012; Jack & Suri, 2011). However, this stream of literature has not yet considered one of the most prominent digital financial services, that is, digital credit. We show that the gender gap in formal digital credit in contrast to other digital financial services is substantial and persistent over time. The present study is also the first to investigate which socio-economic variables explain the persistent gender gap and thereby carves out explicit policy recommendations.

Third, by comparing credit usage in the regulated formal and unregulated semi-formal digital credit markets, we also add to the literature investigating regulation and competition in credit markets (Cuesta & Sepulveda, 2018; Jaffee & Stiglitz, 1990). The present study is the first to compare the association of socio-economic variables between the formal and semi-formal digital credit markets. We show that socio-economic gender differences, in line with economic theory, do not translate into gender differences in digital credit usage in the unregulated semi-formal credit market. We thereby add to thriving policy discussions, which have recently contributed to the loosening of regulations in both formal credit markets, including the retraction of interest rate caps in all credit markets in Kenya.

For policymakers, the present study underscores the long-standing recommendation to decrease women's unfavourable position in Kenya, mainly with regard to income, education, and mobile phone ownership. It also highlights that the recent measure of the Kenyan government to retract interest rate caps might prove particularly fruitful for the financial inclusion of women in the short run.

We proceed by reviewing relevant literature about digital credit. We then describe the data and methods used in the present study. Next, we analyse gender differences in the semi-formal and formal digital credit market in Kenya. Finally, we derive policy implications and conclude.

2 | DIGITAL CREDIT BACKGROUND

Three attributes are typically used to distinguish digital from conventional credit (Chen & Mazer, 2016). First, digital credit is instant; this means that the time-span between application and approval is shorter than 24 h. Second, digital

credit is automated; this refers to the automated credit approval process, which is undertaken by algorithms, based on non-traditional data (such as mobile money or phone usage data). Due to the second attribute of digital credit, borrowers typically do not need to provide collateral. This could be particularly beneficial for women, as women often have lesser control over resources accepted as collateral (Fletschner & Kenney, 2014). Third, digital credit is remote; this refers to the possibility of conducting the entire process from credit application to credit approval from anywhere as long as the borrower has access to mobile network. This attribute may also be particularly useful for women, who often have reproductive roles in the household, not allowing them to leave the house for a long time period. This circumstance is shown to render long travel times to banks a relatively larger obstacle for women (Baydas et al., 1994; Lycette & White, 1989; Moser, 1993). Due to the remote nature of digital credit, this constraint becomes largely obsolete.

Other than the universal approach distinguishing digital credit from conventional credit, the status quo of digital credit is defined by rather short-term, expensive, low-sized loans (Francis et al., 2017). From an economic perspective, these features can help hedge the risk of default (Hoff & Stiglitz, 1990). This enables digital lenders to offer credit to a larger share of the population compared to formal conventional credit lenders. It follows that through formal digital credit a larger share of the population could become able to use a form of formal credit and start to build a financial history (Blechman, 2016; Brailovskaya et al., 2021). However, not all digital lenders can (or are willing to) adjust contract terms according to the risk assessment of the applicant.

Formal digital lenders in Kenya (i.e., where at least one of the lending parties is a bank) are bound to the regulatory framework of the Central Bank Kenya (Blechman, 2016). While this regulatory framework is meant to protect consumers from excessively high fees, it can also cause credit rationing when the set of available contract terms does not allow lenders to make credit offers to individuals whose risk assessment is above a certain threshold (Hoff & Stiglitz, 1990). Especially poorer individuals who are less active users of mobile phones might not generate enough data to be scored by algorithms or may be assessed as too risky given the available set of contract terms (Blechman, 2016). In fact, recent studies show that women use mobile phones less often than men and that women are assigned with lower digital credit scores on average (Brailovskaya et al., 2021). In such cases, regulation might impede financial inclusion through formal digital credit. This may be one reason why all interest rate caps in Kenya have been retracted recently (this policy measure took place after the latest survey round in the present study). However, it should be noted that the formal digital credit market in Kenya is largely dominated by a single credit product, M-Shwari, which is offered at a fixed interest rate only (Suri et al., 2021). Whether the retraction of interest rate caps therefore indeed leads to a large heterogeneity in available contract terms remains dubitable.

In contrast to formal lenders, semi-formal digital lenders in Kenya are not bound to the same set of regulations and can set interest rates arbitrarily high (Blechman, 2016); borrowers of semi-formal digital credit are therefore not considered financially included¹ (Demirgüç-Kunt & Klapper, 2013). Being able to set interest rates arbitrarily high might enable semi-formal lenders to grant credit to individuals which are considered too risky for formal digital lenders; such as women whose lower average credit scores might lead to being disproportionately often considered too risky in the formal digital credit market (Blechman, 2016; Brailovskaya et al., 2021). The downside of accepting riskier borrowers is that default rates are expected to increase. Indeed, due to comparatively high default rates and questionable loan enforcement practices, unregulated semi-formal digital lenders in Kenya are by now prohibited from reporting repayment behaviour to credit reference bureaus (CBK, 2020). While this restriction for semi-formal lenders might protect consumers, it also underscores that those borrowing in the semi-formal digital sector exclusively cannot build up credit scores, and might remain financially excluded. At the same time, it should be noted that the number of semi-formal digital lenders is much higher compared to their formal counterpart. This may be attributable to the lack of regulation in this market, which causes low entry barriers. In 2017, more than 110 semi-formal digital loan products were available in Kenya (Gwer et al., 2019). This competition has thereby led to a much greater variety in available contract terms in that market; including those with excessively high interest rates. From an economic perspective the large variety of contract terms increases the chances of people to find a credit product which they qualify for.

3 | DATA AND METHODS

3.1 | Data

The present study is based on the 2009, 2012, 2015 and 2018 FinAccess household surveys. The surveys represent the second to fifth in a series of surveys, which aim to gather information about the usage of financial services in Kenya. The FinAccess household surveys emanate from a collaboration between the Kenya National Bureau of Statistics (KNBS), Financial Sector Deepening Trust (FSD) Kenya and the Central Bank of Kenya (CBK). The surveys employ a two-stage stratified cluster sampling design to collect data. The first level encompasses the selection of clusters based on the National Sample Surveys and Evaluation Program, with no substitute households being allowed. In the second level, households are selected in each cluster. From each chosen household, an individual aged 16 or above is then randomly selected. According to the FSD, this procedure ensures statistically valid and nationally representative samples for individuals aged 16 and above. We excluded all individuals who are less than 18 years of age from the respective survey rounds as formal credit lenders require borrowers to be 18 years of age or older.

The 2009 FinAccess household survey was conducted in February 2009 and resulted in 6598 respondents; of which three minors were excluded for the present analysis (FSD Kenya, CBK, & KNBS, 2009). The 2012 FinAccess household survey was conducted during the time period of October 2012 to February 2013 and resulted in 6449 respondents, of which 263 minors were excluded for the present analysis (FSD Kenya, CBK, & KNBS, 2013). The 2015 FinAccess household survey was conducted during the time period August to October 2015 and resulted in 8865 respondents, of which 457 minors were excluded for the present analysis (CBK, FSD Kenya, & KNBS, 2016). Finally, the 2018 FinAccess household survey was conducted during the time period of October to December 2018 and resulted in 8669 respondents, of which 402 minors were excluded for the present analysis (FSD Kenya, CBK, & KNBS, 2019). The latest survey contains a unique set of questions regarding borrowing behaviour. Our analysis of why potential differences between women and men with regard to digital credit usage exist therefore focuses on this particular survey. For this in depths analysis, we also excluded all individuals with missing values from the present analysis, leading to a total sample of 7583 individuals. We further conducted appropriate weighting adjustments for all survey rounds in order to provide results which are nationally representative.

3.2 | Methods

The first objective of the present study is to investigate whether formal digital credit decreases the gender gap in financial inclusion in Kenya through total formal credit usage over time. In order to achieve that aim, we compare how total formal credit usage, consisting of either formal digital credit and/or formal conventional credit (i.e., the lending institution is a bank or the government) usage, has changed for men and women between 2009 and 2018. We use different definitions of credit usage to answer the first and second research objectives. To answer the first research objective, we define credit usage (conventional and digital) as having ever used the respective credit form. This definition is chosen for two reasons: First, this definition is suitable for comparisons over time, and second, all survey rounds used the exact same question to elicit whether a responded has ever used the respective credit forms. The latter ensures that the presented changes in credit usage are accurately reflecting real changes in the population and are not due to changes in the framing of the question. To answer our research question, we first calculate whether gender differences in the respective years are statistically significant (first differences). Based on the first differences we then analyse changes in gender differences over time (second differences).

The second objective of the present study is to understand which factors explain differences in formal digital credit usage between men and women in Kenya; specifically, with respect to gender differences in socio-economics as well as the role of market characteristics in the formal digital credit market. In order to achieve these aims we

estimate four different models. Models 1 and 2 include formal digital credit usage as an outcome variable; these models indicate how socio-economic characteristics are associated with formal digital credit usage. Model 3 and 4 include semi-formal digital credit usage as an outcome variable and are estimated to understand whether and why gender differences might not persist in the unregulated semi-formal digital credit market. As socio-economics are not stable over time (e.g., income and age), asking whether an individual has ever used the respective form of credit might be inappropriate. The 2018 FinAccess survey entails an additional question, which elicits whether individuals have used the respective credit forms in the past 12 months. As this more appropriate definition is only available in the 2018 FinAccess Survey, we answer the second research question based on this specific survey only.

The binary nature of the outcome variables implies that a logit model can be used to investigate the relationship between digital credit usage and the set of explanatory variables (Hoetker, 2007). Previous studies which investigate gender gaps have also relied on propensity score matchings and Blinder-Oaxaca decomposition (Bauer et al., 2007; Fonseca et al., 2012; Robinson-Cimpian et al., 2014). However, the functional form assumptions of the parametric Blinder Oaxaca can lead to misleading results (Mora, 2008). The propensity score matching may hence be more appropriate. However, the PSM is particularly suitable when seven or fewer events per confounder exist; with more confounders, logistic regression should be the technique of choice (Cepeda et al., 2003). Given that there are roughly 630 observations per confounder on average in the present study, logistic regressions are the most appropriate estimation technique.

In the first step we estimate the Models 1 and 3, which can be expressed as

$$y_i = \alpha + \beta \text{Female}_i + \mu \quad (1)$$

where y_i is a binary variable that equals one if an individual has used a formal digital credit (Model 1) or semi-formal digital credit (Model 3) within the past 12 months prior to the survey and zero otherwise. Female_i is a binary variable that takes value one if the respondent is female and zero otherwise. The random error term μ is expected to have a standard logistic distribution; α and β are parameters to be estimated in Equation (1).

In order to investigate which confounders explain gender differences in digital credit usage, we estimate Models 2 and 4, which can be expressed as

$$y_i = \alpha + \beta \text{Female}_i + \gamma X_i + \mu \quad (2)$$

where y_i is a binary variable that equals one if an individual has used a formal digital credit (Model 2) or semi-formal digital credit (Model 4) within the past 12 months prior to the survey and zero otherwise. Female_i describes the same variable as in Equation (1). The random error term μ is expected to have a standard logistic distribution; α , β , and γ are parameters to be estimated in Equation (2). X_i is a vector of control variables, including socio-demographic and socio-economic variables. Even though some previous studies suggest to include all clinically and intuitively relevant variables in the model regardless of their statistical significance, such a procedure can introduce numerically unstable estimates in the model and difficulties for other researchers to reproduce the results (Hosmer et al., 2013). Due to these shortcomings, we follow seven steps to purposefully select the model (cf. Hosmer et al., 2013, for an in depths description).

Based on previous studies in the field of credit access and MFS adoption, we initially select 12 control variables as potentially relevant, including age of the respondent and age squared (age^2) to capture non-linear effects (Demombynes & Thegeya, 2012; Munyegera & Matsumoto, 2016); distance to nearest financial institution (Munyegera & Matsumoto, 2016); education level of the respondent (Munyegera & Matsumoto, 2016); whether the respondent is the household head (Tran et al., 2016); the respondent's income, transformed into log form (Demombynes & Thegeya, 2012); a dummy for whether the respondent is married (Demombynes & Thegeya, 2012); a variable that captures whether the respondent owns a mobile phone (Munyegera & Matsumoto, 2016) and/or a national ID (Gwer et al., 2019); the number of household members (Munyegera & Matsumoto, 2016); a variable

which captures province fixed effects (Sekabira & Qaim, 2017); and a dummy that captures whether a respondent lives in an urban or a rural environment (Weil et al., 2012). During the purposeful selection, four of the 12 control variables that have been initially selected are excluded due to their statistical insignificance and absence of adjustment effects; that is, marital status, household head, urban and number of household members are excluded. We also conduct robustness checks including all 12 control variables. In addition, of all clinically or intuitively relevant interaction terms only gender by education is statistically significant albeit only across a fraction of the possible combinations. The interaction term is not included in the main model but discussed separately.

3.3 | Model specification

We first examine whether the analysis is subject to multicollinearity, a common identification problem with cross-sectional data, by generating a pairwise correlation matrix (Table A1 in Appendix A). The matrix shows the degree of correlation amongst the variables. The highest absolute correlation coefficient is -0.34 , between “education” and “age.” We further estimate the variance inflation factor (VIF) between the variables. For both models in which we add control variables (i.e., Models 2 and 4) VIFs are between 1.06 and 2.88 with a mean VIF of 1.92. We conclude therefore that multicollinearity does not seem to cause adverse consequences to the model.

The goodness of fit of the model is then tested with the Archer and Lemeshow test. This test was developed by Archer and Lemeshow (2006); it is similar to the more known Hosmer and Lemeshow test, but contrarily, accounts for the complex design of the survey sampling (Archer & Lemeshow, 2006). We find high p values of 58.9 and 14.1 for both models 2 and 4 respectively, strongly indicating no poor fit of the respective models. To assess whether all the relevant explanatory variables are included in the model, we run a link test. The link test uses the predicted value and its square as the predictors to rebuild the model. The predictor variable should be statistically significant, but not its square. If the latter is statistically significant, this could mean the functional form of the model is inadequate or that relevant variables are omitted (Dudek & Lisicka, 2013). In both models the predictor variable (\hat{y}) is statistically significant but not its square (\hat{y}^2), strongly indicating that the functional form of the respective models is not inadequate.

4 | RESULTS AND DISCUSSION

4.1 | Descriptive statistics of borrower characteristics

Overall, 51.5% of individuals in the 2018 FinAccess Survey are women, and 48.5% are men (Table 1). This share of women is close to the national average in Kenya, of 50.5% (KNBS, 2019).

As a pre-condition to be able to use any form of digital credit, individuals need a valid (national) ID. The vast majority of respondents do fulfil this pre-condition, with a mean of 94.4% owning a national ID. Differences between men and women in national ID ownership are neglectable, contrasting prior research that attributed women's lower access to financial services to lower levels of ID ownership (Demirgüç-Kunt et al., 2013). Another pre-condition to be able to use any form digital credit is having a mobile phone; mobile phone ownership among respondents is also high at 83.9%. Here, we find statistically significant differences between genders, with men being 5.4 percentage points more likely than women to own a mobile phone. We further find that the majority of individuals live less than 10 min walking distance from the nearest financial institution (60.0%). A circumstance that can be largely attributed to the vast number of mobile money agents, more than 200 000 in 2018 (CBK, 2021).

Table 1 further corroborates previous findings on women's unfavourable position in the Kenyan society, showing that men's average income is statistically significantly higher, with a difference of more than 50%. In addition, women are statistically significantly more likely to not have any form of formal education, with a considerable

TABLE 1 Socio-economic descriptive statistics of respondents (in the year 2018)

Borrower characteristics	Explanation	Total Sample ^a Mean (SD)	Men ^a Mean (SD)	Women ^a Mean (SD)	Difference women/men ^b
Observations		7583	3172	4411	
Gender (%)	Dummy (takes value 1 if respondent is female and zero otherwise)	51.54	0	100	
Infrastructure access					
Financial institution Distance	Walking Distance to nearest financial institution in minutes				
<10 min		60.02	61.33	58.81	-2.52*
10–30 min		28.20	27.91	28.48	0.58
>30 min		11.77	10.76	12.72	1.96**
Mobile phone ownership (%)	Dummy (takes value 1 if respondent owns a mobile phone and zero otherwise)	83.93	86.91	81.54	-5.37***
National ID ownership (%)	Dummy (takes value 1 if respondent owns a national ID and zero otherwise)	94.35	94.04	94.63	0.59
Socio-demographics					
Age (years)	Age in years	39.97 (16.39)	40.52 (16.03)	39.45 (16.70)	-1.07**
Income (KSh)	Income in Kenyan Shilling	11 082.66 (20 973.58)	13 591.18 (23 977.63)	8723.80 (17 365.36)	-4867.38***
Education (%)					
No education		12.21	7.80	16.35	8.55***
Some primary		22.94	20.86	24.90	4.04***
Primary completed		21.80	22.34	21.29	1.05
Some secondary		11.01	11.84	10.22	1.64*
Secondary completed		18.68	22.20	15.37	-6.83***
Technical/university		13.37	14.96	11.87	-3.09***

Note: Authors' own calculations based on FinAccess survey 2018.

^aAll means are calculated under consideration of the complex survey design using Stata's svy function.

^bStatistical difference is calculated with the adjusted Wald test.

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

difference of 8.6 percentage points. Men in contrast are statistically significantly more likely to have completed secondary education or even continued education thereafter, with considerable differences of 6.8 percentage points and 3.1 percentage points, respectively.

4.2 | Formal digital credit and financial inclusion through credit usage

In this section, we investigate the influence of formal digital credit on the gender gap in the financial inclusion through total formal credit usage. As described in the data and methods section, total formal credit usage (digital and conventional) is defined in this section as having ever used the respective credit form.

Table 2 shows that between 2012 (shortly before digital credit was introduced in Kenya) and 2018, the gender gap in financial inclusion through total formal credit usage has almost doubled from 4.1 percentage points in 2012 to 7.8 percentage points in 2018, a statistically significant increase (at the 5% level) of 3.7 percentage points. In the same time period, the gender difference in the conventional credit market has only marginally and not statistically significantly increased by 0.7 percentage points. This shows that the considerable increase in the gender gap in total formal credit usage is largely driven by formal digital credit, where gender differences since digital credit's market launch remain high at almost 6 percentage points. More precisely, the results imply that gender differences in formal digital credit usage account for at least 3 percentage points of the total 3.7 percentage point increase in the gender gap in total formal credit usage between 2012 and 2018. This finding strongly indicates that formal digital credit has led to an increase (rather than the hoped-for decrease) in the gender gap in financial inclusion. Of course, the counterfactual scenario in which formal digital credit is not launched remains unknown. Consequently, it could have been that the gender gap in formal conventional credit would have increased much more if formal digital credit had not come to exist. This potential counterfactual scenario however seems unlikely as the two formal credit markets are largely mutually exclusive and complementary in nature; meaning that individuals who borrow in the digital credit market are not doing so instead of borrowing in the conventional market (Suri et al., 2021). It therefore seems more plausible that digital credit has at least partially caused an increase in the gender gap in total formal credit usage.

Despite the finding that formal digital credit appears to have increased the gender gap in financial inclusion, it should be noted that the respective shares of both men and women who are financially included through the usage of total formal credit have increased statistically significantly since the introduction of digital credit in 2012. Interestingly, the increase for men took place especially between 2012 and 2015 (+12.4 percentage points), while between 2015 and 2018 the share of men who are financially included through total formal credit usage has only marginally and not statistically significantly changed (−0.4 percentage points). In contrast, the share of women, who are financially included through total formal credit usage has statistically significantly increased in both periods. However, as for men, the bulk of the increase took place between 2012 and 2015 after formal digital credit was introduced (+6.5 percentage points); while the increase between 2015 and 2018 was much smaller (+1.8 percentage points).

Table 2 further shows that in the period after the introduction of formal digital credit the gender gap increased statistically significantly by more than 240% to 9.9 percentage points. In the most recent period, the gender gap then reduced by 2.1 percentage points, albeit this change is not statistically significant. Table 2 further shows that the decrease in the gender gap is almost entirely driven by changes in the conventional market. In contrast, the gender gap in formal digital credit usage has merely decreased by 0.3 percentage points and remains statistically significant at almost 6 percentage points. As a consequence, the gender gap in the formal digital credit market, despite the lack of requirements for collateral or long-travel times to banks in this market, has become larger than the gender gap in the formal conventional credit market. In the next section, we therefore analyse which reasons could explain the persisting gender gap in the formal digital credit market.

4.3 | Econometric results

Estimation results from logistic regressions of gender on the likelihood to use formal digital credit (Models 1 and 2) and semi-formal digital credit (Models 3 and 4) are presented in Table 3 (all estimations are calculated using STATA 16). Credit usage in this section is defined as having used any of the two respective credit forms within 12 months prior to the survey. This definition is more appropriate (compared to having ever used the respective credit form) for

the following analysis as not all confounders included in the models are likely to be stable over a long time period. This more appropriate definition of digital credit usage is only available for the 2018 FinAccess survey. This section is therefore based on the 2018 FinAccess survey solely. It further follows that due to the adjusted definition of credit usage, credit usage rates in this section are slightly different compared to the previous section. In order to facilitate the interpretation of the results, we also report marginal effects. The marginal effect represents the effect a one-unit change of the explanatory variable has on the dependent variable, while holding all other variables constant (Karaca-Mandic et al., 2012).

4.3.1 | Formal digital credit usage

We find that on average the share of women using formal digital credit is 5.1 percentage points smaller compared to men (Model 1). Controlling for other factors reduces the association between gender and formal digital credit usage considerably, albeit remaining marginally statistically significant; with women being 2.2 percentage points less likely compared to men to use formal digital credit controlling for other factors (Model 2). We interpret this finding as strong indication that differences in gender with regard to the observed control variables, rather than gender itself, largely explain women's lower likelihood to use formal digital credit. The finding is also congruent with previous literature stating that women are more likely to be credit-constrained compared with men of equivalent socio-economic characteristics in other credit markets (Diagne et al., 2000; Fletschner, 2009).

We further show that, as expected, mobile phone ownership is statistically significantly and considerably associated with formal digital credit usage. As seen in Table 1, women are statistically significantly less likely to own a mobile phone; where differences are also large in magnitude at 4.7 percentage points. The absence of statistically significant interaction effects between gender and mobile phone ownership further implies that a decrease in the gender gap in mobile phone ownership is expected to be associated with a decrease in gender related differences in formal digital credit usage. In addition to mobile phone ownership, another pre-condition to use digital credit, is the ownership of a (national) ID, which is congruently statistically significantly and considerably associated with formal digital credit usage. A third variable, that can be subsumed under infrastructure to access formal digital credit usage, is the proximity to the nearest financial institution (most typically mobile money agents). Congruently to previous studies, an increase in the distance to financial providers is negatively (and statistically significantly, when comparing individuals who live less than 10 min and those who live between 10 and 30 min from the nearest institution) associated with the probability that an individual uses formal digital credit; most likely as transaction costs increase with distance (Jack & Suri, 2014).

We also find that an increase in age by 1 year is statistically significantly positively related to the probability of formal digital credit usage, while the negative and statistically significant (albeit small) association of age² and formal digital credit usage indicates a non-linear, decreasing relationship with increasing age. These results are also in line with research in the field of credit access (Freeman et al., 1998; Jia et al., 2010; Loges & Jung, 2001). Further, Model 2 shows that an increase in education level is positively associated with the probability that an individual uses formal digital credit; this finding is consistent for both genders across all educational levels (not shown). It follows that a decrease in gender differences with regard to education would also translate in decreasing gender differences in formal digital credit usage. Nevertheless, it should be noted that while an increase in education at all levels is associated with a statistically significant increased probability to use digital credit for women, across some educational levels gender related interaction effects occur (Table A2). Calculating interaction effects in non-linear models is not as straightforward as in linear models and requires the calculation of the second differences across all possible groups (i.e., six educational levels times two genders yields 12 groups). Four of the 12 possible constellations show statistically significantly higher increases for men in the likelihood to use digital credit when education increases (Table A2). However, the main finding that a decrease in education related gender differences is associated with a decrease in gender differences in formal digital credit usage still holds.

TABLE 2 Formal credit usage rates by gender over time

Men		Women		Comparison Men and Women	
Share of users (%)	Change in share of users (percentage points)	Share of users (%)	Change in share of users (percentage points)	Gender differences in the share of users (percentage points)	Change in gender differences in the share of users (percentage points)
Formal credit usage (total share)					
2009	5.45 (4.72; 6.29)	2.59 (1.72; 3.90)	n.a.	2.9***	n.a.
2012	11.15 (9.35; 13.26)	7.10 (6.08; 8.25)	+5.7***	4.1***	+1.2
2015	23.50 (21.77; 25.33)	13.56 (12.46; 14.73)	+12.4***	9.9***	+5.9***
2018	23.18 (21.10; 25.29)	15.37 (13.86; 17.02)	-0.04	7.8***	-2.1
Formal conventional					
2009	5.45 (4.72; 6.29)	2.59 (1.72; 3.90)	n.a.	2.9***	n.a.
2012	11.15 (9.35; 13.26)	7.10 (6.08; 8.25)	+5.7***	4.1***	+1.2
2015	14.24 (12.94; 15.66)	7.78 (7.02; 8.61)	+3.1***	6.5***	+2.4*
2018	11.35 (9.89; 12.98)	6.05 (5.65; 7.57)	-2.9***	4.8***	-1.7
Formal digital					
2009	0	0	n.a.	n.a.	n.a.
2012	0	0	n.a.	n.a.	n.a.
2015	13.51 (12.21; 14.92)	7.44 (6.62; 8.35)	+13.5	6.1***	n.a.
2018	16.50 (14.76; 18.39)	10.66 (9.38; 12.10)	+3.0***	5.8***	-0.3

Note: Formal credit usage indicates whether the respondent has ever used formal conventional and/or formal digital credit. All shares are calculated under consideration of the complex survey design using Stata's svy function. Authors' own calculations based on the FinAccess household survey 2009, 2012, 2015, and 2018; N (Men): 4971 (2009); 2517 (2012); 3181 (2015); 3295 (2018); N (Women): 1624 (2009); 3669 (2012); 5027 (2015); 4567 (2018); 95% confidence intervals are shown in parentheses. Statistical difference is calculated with the adjusted Wald test.

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

Lastly, Table 3 shows that increases in income are statistically significantly associated with increases in the likelihood to use formal digital credit for both men and women. The result is also in line with previous studies in credit access literature that find a positive and statistically significant relationship between income and credit access (Feder et al., 1990). It further corroborates that algorithms might assess especially poorer individuals as too risky, given the available set of contracts in the regulatory framework (Blechman, 2016). In the next section, we therefore compare the presented results to the semi-formal digital credit market, where the set of available contract terms is considerably larger.

4.3.2 | Semi-formal digital credit usage

Table 3 shows that there are no statistically significant gender differences in the semi-formal, unregulated digital credit differences (Models 3 and 4 in Table 3). As described in the section on digital credit background, semi-formal lenders are not bound to the same regulatory framework as formal lenders, increasing the set of contract terms they can offer and which might decrease market entry barriers (Blechman, 2016; Hoff & Stiglitz, 1990). This has led to a large variety of semi-formal digital lenders in Kenya who offer digital credit with a wide array of contract terms, including excessively high interest rates; with more than 110 semi-formal digital loan products which were available in Kenya in 2017 (Gwer et al., 2019). From an economic viewpoint, this in turn allows semi-formal lenders to serve loans to riskier borrowers (Hoff & Stiglitz, 1990). In line with this theoretical expectation, we find that variables which are often used to assess the riskiness of borrowers defaulting (Agarwal et al., 2011; Atieno, 2001), foremost income size, are not statistically significantly associated with the usage of semi-formal digital credit, contrasting to formal digital credit. We cautiously interpret this finding as indication that market characteristics, foremost with regard to available contract terms, might partially moderate the association between socio-economic characteristics and digital credit usage. More precisely, we interpret our findings as suggestive evidence that the socio-economic gender differences in combination with the specific market characteristics of the formal digital credit market (especially the lack of heterogeneity in available contract terms) explain the gender gap in formal digital credit to a large extent.

4.4 | Robustness and model specification

We run multiple robustness checks to validate our main findings. First, we conduct several specifications of the logit models by including variables, which were removed during purposeful model selection; including the interaction between female and education; including individuals of all ages and imputing income (as these two variables largely account for the excluded individuals). Second, we estimate a linear probability model (LPM), which delivers unbiased and consistent effects for the average effects estimated (Wooldridge, 2010). Third, we analyse which share of the gender gap in formal digital credit usage can be explained by observable variables, employing both a propensity score matching (PSM) and a Blinder-Oaxaca decomposition.

We find very similar estimates for the association between being female and formal digital credit usage for the different specifications of the logit model (Table A2). This finding strongly corroborates that gender differences in socio-economic variables explain the major part of differences in formal digital credit usage. The other coefficients also remain stable in sign, magnitude or statistical significance levels for the different specifications of the logit models, giving further confidence to the robustness of the association between socio-economic variables and digital credit usage.

The estimates of the LPM of Model 2 are also largely similar to the estimates of the other robustness checks in sign and statistical significance (Table A2). The magnitude of the association, in contrast, is considerably lower in the LPM for the association of both owning an ID and owning a mobile phone with formal digital credit usage. These

TABLE 3 Formal and semi-formal digital credit usage and gender (N = 7583)

Variables	Formal		Semi-formal	
	Model (1)	Model (2)	Model (3)	Model (4)
	Logit coefficients	Marginal effect	Logit coefficients	Marginal effect
Independent variable				
Female (1 = yes)	-0.543*** (0.092)	-0.270*** (0.099)	-0.022*** (0.008)	-0.004 (0.007)
			-0.060 (0.100)	-0.009 (0.106)
Infrastructure access				
Mobile phone ownership (1 = yes)	2.552*** (0.717)	0.208*** (0.058)	1.158*** (0.269)	0.076*** (0.018)
National ID ownership (1 = yes)	1.520** (0.600)	0.124** (0.049)	0.724** (0.325)	0.048** (0.024)
Financial institution distance (Base = <10 min)				
10–30 min	-0.338** (0.143)	-0.026** (0.010)	-0.434*** (0.139)	-0.027*** (0.008)
>30 min	-0.205 (0.212)	-0.017 (0.016)	-0.182 (0.206)	-0.012 (0.013)
Socio-demographics				
Age (years)	0.066*** (0.023)	0.005*** (0.002)	-0.012 (0.023)	-0.001 (0.002)
Age ²	-0.001*** (0.000)	(0.000)	-0.000 (0.000)	(0.000)
Income (ln of KSh)	0.329*** (0.059)	0.027*** (0.005)	0.027 (0.042)	0.002 (0.003)
Education (Base = No education)				
Some primary	1.505** (0.587)	0.033*** (0.009)	0.859** (0.336)	0.031*** (0.010)
Primary completed	2.249*** (0.575)	0.073*** (0.010)	1.107*** (0.336)	0.045*** (0.010)

TABLE 3 (Continued)

	Formal		Semi-formal	
	Model (1)	Model (2)	Model (3)	Model (4)
Some secondary		2.804*** (0.584)	0.122*** (0.016)	1.013*** (0.365)
Secondary completed		2.726*** (0.569)	0.114*** (0.011)	1.425*** (0.343)
Technical/university		3.151*** (0.575)	0.162*** (0.014)	1.763*** (0.346)
Province				
North Eastern/Coast		-0.693*** (0.231)	-0.060*** (0.019)	0.476** (0.213)
Eastern		-1.278*** (0.211)	-0.092*** (0.015)	0.864*** (0.199)
Central		-0.005 (0.187)	-0.001 (0.019)	1.353*** (0.216)
Rift Valley		-0.497*** (0.160)	-0.045*** (0.015)	-0.686*** (0.222)
Western/Nyanza		-0.437*** (0.168)	-0.040** (0.016)	0.209 (0.205)
Constant		-11.919*** (1.142)		-5.112*** (0.667)
hat (p value)		0.000		0.000
hat ² (p value)		0.396		0.652
Archer-Lemeshow Test		0.589		0.1405

Note: Credit usage defined as credit usage in the past 12 months. Province fixed effects and Province × Financial Institution Distance interactions are only shown in Appendix A due to the large number of variables (Tables A1 and A2). Robust standard errors are shown in parentheses. Authors' own calculations based on 2018 FinAccess Household Survey.
* $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

deviations in coefficient sizes between logit models and LPMs could for instance be due to unobserved heterogeneity which is uncorrelated with the confounders, which can alter the magnitude of the confounders in the logit model, but not in the LPM.

Table A3 shows that the association between gender, income and semi-formal digital credit usage in Model 4 (Table 3), is robust in sign, magnitude and statistical significance throughout the different logit model specifications. The LPM further indicates that changes in education level do not statistically significantly increase the likelihood that people use semi-formal digital credit, unless they have completed secondary education or higher.

As the present study only investigates average effect estimates, the LPM analysis delivers unbiased and consistent estimators (Wooldridge, 2010). We can therefore estimate whether omitted variable bias could disprove the main findings, by using the LPMs. We estimate robustness values (RVs) for the estimates of “female” (4.61%), “mobile phone ownership” (3.21%), “education” (15.10%) and “income” (9.26%) in Model 2. The RVs indicate a threshold of how much of the residual variance, both of the treatment and the outcome, an unobserved confounder would need to explain to bring the lower bound of the confidence interval (of the estimated effect) to zero (significance level 5%) (Cinelli & Hazlett, 2020). To facilitate the interpretation of what that means, we use income as a benchmark for the former three RVs (i.e., for female, mobile phone ownership, and education), as income is correlated with all confounders and a strong predictor of digital credit usage. We find in each case that unobserved confounders as strong as income would not be sufficient to explain the observed estimate; with the higher bound estimates for income being below the respective RV in each case, at 1.51% (cf. RV female), 1.44% (cf. RV mobile phone ownership) and 3.71% (cf. RV education). To interpret the RV of income, we use mobile phone ownership as benchmark. We find that an unobserved confounder as strong as mobile phone ownership (higher bound at 1.44%) would not be sufficient to explain the observed estimate on income (cf. RV income).

A main finding of the second research question is that socio-economic differences explain the major share of gender differences in formal digital credit usage. We shortly discussed in the section data and methods, why logit models are an appropriate choice for the present analysis. Nevertheless, we compare the main finding of the second research question to two other statistical methods, which are often used to analyse gender gaps (i.e., Propensity Score Matching and Blinder-Oaxaca decomposition). First, we estimate a propensity score model (PSM). This method enables us to compare men and women with similar covariate values in the observed characteristics, so that comparisons between genders are not confounded by differences in covariate distributions (Rubin, 1997). We choose nearest neighbour matching as a matching technique as this form of matching has been shown to perform consistently well across a wide range of settings (Austin & Stuart, 2017). In order to obtain excellent balance for all covariates we group income in ten groups; the other covariates are the same as in Model 2. The PSM estimates an average treatment effect on the treated of 1.4 percentage points, which is marginally statistically significant at the 10% level (Table A4). In other words, being female decreases the probability of formal digital credit take up by 1.4 percentage points, when compared to men with similar propensity scores. Second, we conduct a Blinder-Oaxaca decomposition, which is often used to decompose average gender gaps into an explained and unexplained part (Bauer et al., 2007; Blinder, 1973). The results indicate that 3.1 percentage points of the 5.1 percentage point gender gap can be explained by differences in the observed covariates (Table A5). This can be interpreted as 61.8% of the gender gap in formal digital credit usage being explained by gender differences in socio-economic variables. We view the findings from the last two robustness checks as strong support that the bulk of the gender gap in formal digital credit usage can be explained by differences in the observable socio-economic variables.

5 | CONCLUSION AND RECOMMENDATIONS

The present study reveals that formal digital credit has largely contributed to a statistically significant increase in the gender gap in financial inclusion. The present study therefore clearly shows that despite formal digital credit characteristics, which seem particularly beneficial for women (such as the lack of collateral requirements), as of

now formal digital credit is no means to decrease the gender gap in financial inclusion; but rather a pathway towards increasing it. The study yields further insights as to why the gender gap in formal digital credit usage in Kenya persists. We find strong indication that gender differences in socio-economic variables (foremost in education, income, and mobile phone ownership) largely explain differences in formal digital credit usage. We further find that socio-economic gender differences do not lead to a gender gap in the unregulated semi-formal digital credit market. We suggest, in accordance with economic literature, that the absence of a gender gap in the semi-formal digital credit market could be partially due to the larger heterogeneity in available contract terms compared to its formal counterpart.

The analysis therefore delivers valuable indications for policymakers, who aim to decrease the gender gap in financial inclusion. First, we find strong indication that strengthening women's position in society, foremost with respect to education, income and mobile phone ownership, is likely to translate into narrowing the gender gap in the formal digital credit market. Second, we find some indication that fostering a greater variety of formal digital contract terms (for instance by removing interest rate caps or stimulating competition) could increase the user share of women in the formal digital credit market. In fact, based on a similar reasoning, the Kenyan government has recently retracted interest rate caps for both formal credit markets.

We hence strongly encourage future research to investigate the effects of the retraction of interest rate caps on financial inclusion of women through formal digital credit. In addition, the present study uncovers further potential to investigate why gender differences exist in the regulated formal credit market and not in the unregulated semi-formal digital credit market; especially with regard to contract term availability. Lastly, even though Kenya is a vanguard for digital credit usage, socio-cultural, political and other relevant differences should be considered when results are transferred to other contexts. We therefore encourage future research to investigate the role of formal digital credit for financial inclusion of women in other countries.

ACKNOWLEDGEMENTS

Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

The data of FinAccess survey 2009, 2012, 2015 and 2018 are available in "Harvard Dataverse" with the digital object identifiers <https://doi.org/10.7910/DVN/VWLG4>, <https://doi.org/10.7910/DVN/JSPE9W> and <https://doi.org/10.7910/DVN/QUTLO2>. The data of the FinAccess survey 2018 are available on the website of the Central Bank of Kenya, <https://www.centralbank.go.ke/finaccess/>.

ORCID

Constantin Johnen  <https://orcid.org/0000-0001-5455-8008>

ENDNOTE

¹ Unless, of course, the individual is considered financially included through the usage of other formal financial instruments

REFERENCES

- Adegbite, O. O., & Macheche, C. L. (2020). Bridging the financial inclusion gender gap in smallholder agriculture in Nigeria: An untapped potential for sustainable development. *World Development*, 127, 104755. <https://doi.org/10.1016/j.worlddev.2019.104755>
- Agarwal, S., Chomsisengphet, S., & Liu, C. (2011). Consumer bankruptcy and default: The role of individual social capital. *Journal of Economic Psychology*, 32(4), 632–650. <https://doi.org/10.1016/j.joep.2010.11.007>
- Andersson-Manjang, S. K., & Naghavi, N. (2021). State of the industry report on mobile money.
- Archer, K. J., & Lemeshow, S. (2006). Goodness-of-fit test for a logistic regression model fitted using survey sample data. *The Stata Journal*, 6(1), 97–105. <https://doi.org/10.1177/1536867X0600600106>

- Atieno, R. (2001). *Formal and informal institutions' lending policies and access to credit by small-scale enterprises in Kenya: An empirical assessment*. AERC research paper: Vol. 111. African Economic Research Consortium.
- Austin, P. C., & Stuart, E. A. (2017). Estimating the effect of treatment on binary outcomes using full matching on the propensity score. *Statistical Methods in Medical Research*, 26(6), 2505–2525. <https://doi.org/10.1177/0962280215601134>
- Banerjee, A., Karlan, D., & Zinman, J. (2015). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7(1), 1–21. <https://doi.org/10.1257/app.20140287>
- Bauer, T., Göhlmann, S., & Sinning, M. (2007). Gender differences in smoking behavior. *Health Economics*, 16(9), 895–909. <https://doi.org/10.1002/hec.1259>
- Baydas, M. M., Meyer, R. L., & Aguilera-Alfred, N. (1994). Discrimination against women in formal credit markets: Reality or rhetoric? *World Development*, 22(7), 1073–1082. [https://doi.org/10.1016/0305-750x\(94\)90149-x](https://doi.org/10.1016/0305-750x(94)90149-x)
- Bharadwaj, P., & Suri, T. (2020). Improving financial inclusion through digital savings and credit. *AEA Papers and Proceedings*, 110, 584–588. <https://doi.org/10.1257/pandp.20201084>
- Björkregren, D., & Grissen, D. (2018). The potential of digital credit to bank the poor. *AEA Papers and Proceedings*, 108, 68–71. <https://doi.org/10.1257/pandp.20181032>
- Blechman, J. G. (2016). Mobile credit in Kenya and Tanzania: Emerging regulatory challenges in consumer protection, credit reporting and use of customer transactional data. *The African Journal of Information and Communication*, 2016(17), 61–88. <https://doi.org/10.23962/10539/21628>
- Bliner, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4), 436–455. <https://doi.org/10.2307/144855>
- Braïlovskaya, V., Dupas, P., & Robinson, J. (2021). Digital credit: Filling a hole, or digging a hole? Evidence from Malawi.
- CBK. (2020). Publication of the credit reference bureau regulations, 2020 and additional measures on credit information sharing [Press release]. Nairobi.
- CBK. (2021). Mobile payments. <https://www.centralbank.go.ke/national-payments-system/mobile-payments/>
- CBK, FSD Kenya, & KNBS. (2016). FinAccess household survey 2015. Harvard Dataverse. <https://doi.org/10.7910/DVN/QUTLO2>
- Cepeda, M. S., Boston, R., Farrar, J. T., & Strom, B. L. (2003). Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *American Journal of Epidemiology*, 158(3), 280–287. <https://doi.org/10.1093/aje/kwg115>
- Chen, G., & Mazer, R. (2016). Instant, automated, remote: The key attributes of digital credit. <https://www.cgap.org/blog/instant-automated-remote-key-attributes-digital-credit>
- Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, 82(1), 39–67. <https://doi.org/10.1111/rssb.12348>
- Cuesta, J. I., & Sepulveda, A. (2018). Price regulation in credit markets: A trade-off between consumer protection and credit access. <https://doi.org/10.2139/ssrn.3282910>
- Demirgüç-Kunt, A., & Klapper, L. (2013). Measuring financial inclusion: Explaining variation in use of financial services across and within countries. *Brookings Papers on Economic Activity*, 2013(1), 279–340. <https://doi.org/10.1353/eca.2013.0002>
- Demirgüç-Kunt, A., Klapper, L. F., & Singer, D. (2013). Financial inclusion and legal discrimination against women: Evidence from developing countries. World Bank Policy Research Working Paper (6416). DOI: <https://doi.org/10.1596/1813-9450-6416>
- Demombynes, G., & Thegeya, A. (2012). Kenya's mobile revolution and the promise of mobile savings.
- Diagne, A., Zeller, M., & Sharma, M. (2000). Empirical measurements of households' access to credit and credit constraints in developing countries: Methodological issues and evidence. <https://ageconsearch.umn.edu/record/16465/files/fc000090.pdf>
- Dudek, H., & Lisicka, I. (2013). Determinants of poverty—Binary logit model with interaction terms approach. *Ekonometria*, 3(41), 65–77. <https://www.cceol.com/search/article-detail?id=110743>
- Feder, G., Lau, L. J., Lin, J. Y., & Luo, X. (1990). The relationship between credit and productivity in Chinese agriculture: A microeconomic model of disequilibrium. *American Journal of Agricultural Economics*, 72(5), 1151–1157. <https://doi.org/10.2307/1242524>
- Fletschner, D. (2009). Rural women's access to credit: Market imperfections and intrahousehold dynamics. *World Development*, 37(3), 618–631. <https://doi.org/10.1016/j.worlddev.2008.08.005>
- Fletschner, D., & Kenney, L. (2014). Rural women's access to financial services: Credit, savings, and insurance. In *Gender in agriculture* (pp. 187–208). Springer. https://doi.org/10.1007/978-94-017-8616-4_8
- Fonseca, R., Mullen, K. J., Zamarro, G., & Zissimopoulos, J. (2012). What explains the gender gap in financial literacy? The role of household decision making. *The Journal of Consumer Affairs*, 46(1), 90–106. <https://doi.org/10.1111/j.1745-6606.2011.01221.x>
- Francis, E., Blumenstock, J., & Robinson, J. (2017). Digital credit: A snapshot of the current landscape and open research questions. CEGA White Paper, 1739–1776. https://people.ucsc.edu/~jmrtwo/digital_credit_landscape.pdf

- Freeman, H. A., Ehui, S. K., & Jabbar, M. A. (1998). Credit constraints and smallholder dairy production in the East African highlands: Application of a switching regression model. *Agricultural Economics*, 19(1–2), 33–44. <https://doi.org/10.1111/j.1574-0862.1998.tb00512.x>
- FSD Kenya, CBK, & KNBS. (2009). FinAccess Household Survey 2009.
- FSD Kenya, CBK, & KNBS. (2013). FinAccess household survey 2013. Harvard Dataverse. <https://doi.org/10.7910/DVN/JSPE9W>
- FSD Kenya, CBK, & KNBS. (2019). 2019 FinAccess household survey.
- Ghosh, S., & Vinod, D. (2017). What constrains financial inclusion for women? Evidence from Indian micro data. *World Development*, 92, 60–81. <https://doi.org/10.1016/j.worlddev.2016.11.011>
- Gubbins, P., & Totolo, E. [E.] (2018). Digital credit in Kenya: Evidence from demand side-survey. Kenya.
- Guerin, I. (2006). Women and money: Lessons from Senegal. *Development and Change*, 37(3), 549–570. <https://doi.org/10.1111/j.0012-155X.2006.00490.x>
- Gwer, F., Odero, J., & Totolo, E. [E.]. (2019). *Digital credit audit report: Evaluating the conduct and practice of digital lending in Kenya*. FSD Kenya.
- Hannig, A., & Jansen, S. (2010). Financial inclusion and financial stability: Current policy issues. <https://doi.org/10.2139/ssrn.1729122>
- Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal*, 28(4), 331–343. <https://doi.org/10.1002/smj.582>
- Hoff, K., & Stiglitz, J. E. (1990). Introduction: Imperfect Information and rural credit markets—Puzzles and policy perspectives. *The World Bank Economic Review*, 4(3), 235–250. <https://doi.org/10.1093/wber/4.3.235>
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons. <https://doi.org/10.1002/9781118548387>
- Jack, W., & Suri, T. (2011). Mobile money: The economics of M-PESA. National Bureau of Economic Research Working Paper No. 16721. Advance Online Publication. <https://doi.org/10.3386/w16721>
- Jack, W., & Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review*, 104(1), 183–223. <https://doi.org/10.1257/aer.104.1.183>
- Jaffee, D., & Stiglitz, J. E. (1990). Credit rationing. In *Handbook of monetary economics* (Vol. 2, pp. 837–888). Elsevier. <https://www.sciencedirect.com/science/article/pii/S1573449805800232>, [https://doi.org/10.1016/S1573-4498\(05\)80023-2](https://doi.org/10.1016/S1573-4498(05)80023-2)
- Jia, X., Heidhues, F., & Zeller, M. (2010). Credit rationing of rural households in China. *Agricultural Finance Review*, 70(1), 37–54. <https://doi.org/10.1108/00021461011042620>
- Johnen, C., Parlasca, M., & Müßhoff, O. (2021). Promises and pitfalls of digital credit: Empirical evidence from Kenya. *PLoS ONE*, 16(7), e0255215. <https://doi.org/10.1371/journal.pone.0255215>
- Karaca-Mandic, P., Norton, E. C., & Dowd, B. (2012). Interaction terms in nonlinear models. *Health Services Research*, 47, 255–274. <https://doi.org/10.1111/j.1475-6773.2011.01314.x>
- KNBS. (2019). 2019 Kenya population and housing census results. <https://www.knbs.or.ke/?p=5621>
- Loges, W. E., & Jung, J.-Y. (2001). Exploring the digital divide. *Communication Research*, 28(4), 536–562. <https://doi.org/10.1177/009365001028004007>
- Lycette, M., & White, K. (Eds.) (1989). *Improving women's access to credit in Latin America and the Caribbean: Policy and project recommendations*. Kumarian Press.
- Mora, R. (2008). A nonparametric decomposition of the Mexican American average wage gap. *Journal of Applied Econometrics*, 23(4), 463–485. <https://doi.org/10.1002/jae.1006>
- Moser, C. O. N. (1993). *Gender planning and development: Theory, practice, and training*. Routledge. <https://www.taylorfrancis.com/books/mono/10.4324/9780203411940/gender-planning-development-caroline-moser>
- Munyegera, G. K., & Matsumoto, T. (2016). Mobile money, remittances, and household welfare: Panel evidence from rural Uganda. *World Development*, 79, 127–137. <https://doi.org/10.1016/j.worlddev.2015.11.006>
- Robinson-Cimpian, J. P., Lubienski, S. T., Ganley, C. M., & Copur-Gencturk, Y. (2014). Teachers' perceptions of students' mathematics proficiency may exacerbate early gender gaps in achievement. *Developmental Psychology*, 50(4), 1262–1281. <https://doi.org/10.1037/a0035073>
- Rubin, D. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine*, 127(127), 757–763. https://doi.org/10.7326/0003-4819-127-8_Part_2-199710151-00064
- Sekabira, H., & Qaim, M. (2017). Mobile money, agricultural marketing, and off-farm income in Uganda. *Agricultural Economics*, 48(5), 597–611. <https://doi.org/10.1111/agec.12360>
- Suri, T., Bharadwaj, P., & Jack, W. (2021). Fintech and household resilience to shocks: Evidence from digital loans in Kenya. *Journal of Development Economics*, 153(153), 102697. <https://doi.org/10.1016/j.jdeveco.2021.102697>
- Swamy, V. (2014). Financial inclusion, gender dimension, and economic impact on poor households. *World Development*, 56, 1–15. <https://doi.org/10.1016/j.worlddev.2013.10.019>

- Tran, M. C., Gan, C. E., & Hu, B. (2016). Credit constraints and their impact on farm household welfare. *International Journal of Social Economics*, 43(8), 782–803. <https://doi.org/10.1108/IJSE-11-2014-0243>
- Weil, D., Mbiti, I., & Mwega, F. (2012). The implications of innovations in the financial sector on the conduct of monetary policy in East Africa. Report Submitted to the International Growth Centre Tanzania Country Program. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.259.7015&rep=rep1&type=pdf>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.

How to cite this article: Johnen, C., & Mußhoff, O. (2023). Digital credit and the gender gap in financial inclusion: Empirical evidence from Kenya. *Journal of International Development*, 35(2), 272–295. <https://doi.org/10.1002/jid.3687>

APPENDIX A.

TABLE A1 Correlation matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Age	1.000												
Education	-0.344	1.000											
Female	-0.032	-0.141	1.000										
Financial Literacy	-0.196	0.529	-0.102	1.000									
Financial Provider Distance	0.138	-0.262	0.032	-0.149	1.000								
Household Head	0.305	-0.064	-0.144	-0.037	0.007	1.000							
Household Size	-0.092	-0.151	0.031	-0.100	0.144	-0.180	1.000						
Income	-0.021	0.276	-0.163	0.249	-0.161	0.176	-0.115	1.000					
National ID	0.207	0.018	0.013	-0.021	-0.005	0.208	-0.080	0.198	1.000				
Married	0.064	-0.024	-0.059	-0.007	-0.003	0.102	0.258	0.150	0.137	1.000			
Mobile Phone Ownership	-0.129	0.291	-0.067	0.206	-0.161	0.115	-0.076	0.257	0.219	0.113	1.000		
Province	0.082	-0.075	-0.000	-0.110	0.068	-0.069	0.107	-0.186	-0.049	0.032	-0.092	1.000	
Urban	-0.218	0.318	-0.029	0.222	-0.355	0.031	-0.215	0.232	-0.017	-0.032	0.174	-0.353	1.000

Note: N = 7583. Authors' own calculations based on 2018 FinAccess Household Survey.

TABLE A2 Robustness checks of Model 2 (formal digital credit)

Variables	Main model	(1) LPM	(2) Added variables	(3) Interaction	(4) All ages + Income imputed
Independent variable					
Female (1 = yes)	-0.270*** (0.099)	-0.024*** (0.008)	-0.203* (0.106)	-0.114 (1.042)	-0.286*** (0.093)
Infrastructure access					
Mobile phone ownership (1 = yes)	2.552*** (0.717)	0.027*** (0.005)	2.524*** (0.718)	2.547*** (0.717)	2.562*** (0.641)
National ID ownership (1 = yes)	1.520** (0.600)	0.049*** (0.013)	1.529** (0.606)	1.517** (0.600)	1.896*** (0.536)
Financial institution distance (Base = Under 10 min)					
10–30 min	-0.338** (0.143)	-0.027*** (0.009)	-0.293** (0.146)	-0.339** (0.143)	-0.373*** (0.138)
>30 min	-0.205 (0.212)	-0.015 (0.010)	-0.147 (0.218)	-0.205 (0.211)	-0.250*** (0.212)
Socio-demographics					
Age (years)	0.066*** (0.023)	0.002** (0.001)	0.059** (0.024)	0.067*** (0.023)	0.0755*** (0.020)
Age ²	-0.001*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Income (ln of KSh)	0.329*** (0.059)	0.021*** (0.003)	0.310*** (0.061)	0.328*** (0.059)	0.325*** (0.058)
Education (Base = No education)					
Some primary	1.505** (0.587)	-0.005 (0.007)	1.529*** (0.590)	1.717** (0.776)	1.513*** (0.552)
Primary completed	2.249*** (0.575)	0.024** (0.011)	2.262*** (0.578)	2.360*** (0.761)	2.257*** (0.535)
Some secondary	2.804*** (0.584)	0.078*** (0.016)	2.824*** (0.585)	2.956*** (0.771)	2.663*** (0.545)
Secondary completed	2.726*** (0.569)	0.083*** (0.013)	2.734*** (0.572)	2.753*** (0.750)	2.626*** (0.530)
Technical/university	3.151*** (0.575)	0.171*** (0.018)	3.162*** (0.579)	3.254*** (0.751)	3.121*** (0.537)
Province					
North Eastern/Coast	-0.693*** (0.231)	-0.100*** (0.025)	-0.660*** (0.235)	-0.698*** (0.231)	-0.757*** (0.214)
Eastern	-1.278*** (0.211)	-0.120*** (0.022)	-1.212*** (0.218)	-1.283*** (0.211)	-1.406*** (0.205)
Central	-0.005 (0.187)	-0.020 (0.029)	0.088 (0.199)	-0.012 (0.186)	0.021 (0.174)
Rift Valley	-0.497*** (0.160)	-0.083*** (0.022)	-0.425*** (0.164)	-0.501*** (0.160)	-0.566*** (1.150)
Western/Nyanza	-0.437*** (0.168)	-0.076*** (0.023)	-0.369** (0.176)	-0.439*** (0.168)	-0.497*** (0.157)
Added variables					
Married	Not included	Not included	0.054 (0.122)	Not included	Not included

TABLE A2 (Continued)

Variables	Main model	(1) LPM	(2) Added variables	(3) Interaction	(4) All ages + Income imputed
Household size	Not included	Not included	0.008 (0.030)	Not included	Not included
Urban	Not included	Not included	0.159 (0.125)	Not included	Not included
Household head	Not included	Not included	0.179 (0.124)	Not included	Not included
Interaction					
Female * Education (2nd difference)	Not included	Not included	Not included		Not included
No education vs. Primary				0.036* (0.020)	
No education vs. Some sec. education				0.064** (0.032)	
No education vs. Technical/university				0.065*** (0.025)	
Secondary compl. vs. Technical/university				0.078*** (0.030)	
Constant	-11.918*** (1.142)	-0.141*** (0.037)	-11.947*** (1.170)	-12.010*** (1.230)	-12.316*** (1.025)
Observations	7583	7583	7583	7583	8561

Note: N = 7583. Robust standard errors are shown in parentheses. Authors' own calculations based on 2018 FinAccess Household Survey.

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

TABLE A3 Robustness checks of Model 4 (semi-formal digital credit)

Variables	Main model	(1) LPM	(2) Added variables	(3) Interaction	(4) All ages + Income imputed
Independent variable					
Female (1 = yes)	-0.009 (0.106)	-0.000 (0.007)	0.028 (0.118)	-1.176* (0.613)	-0.098 (0.0966)
Infrastructure access					
Mobile phone ownership (1 = yes)	1.158*** (0.269)	0.031*** (0.007)	1.140*** (0.271)	1.182*** (0.268)	1.521*** (0.279)
National ID ownership (1 = yes)	0.724** (0.325)	0.046*** (0.012)	0.689** (0.322)	0.726** (0.326)	0.425 (0.291)
Financial institution distance (Base = Under 10 min)					
10–30 min	-0.434*** (0.139)	-0.029*** (0.007)	-0.353** (0.143)	-0.423*** (0.139)	-0.473*** (0.132)
>30 min	-0.182 (0.206)	-0.013 (0.011)	-0.065 (0.209)	-0.196 (0.206)	-0.249 (0.207)
Socio-demographics					
Age (years)			0.002 (0.024)		-0.006 (0.021)

(Continues)

TABLE A3 (Continued)

Variables	Main model	(1) LPM	(2) Added variables	(3) Interaction	(4) All ages + Income imputed
	-0.012 (0.023)	-0.003** (0.001)		-0.015 (0.023)	
Age ²	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Income (ln of KSh)	0.027 (0.042)	0.002 (0.003)	0.015 (0.041)	0.032 (0.042)	0.032 (0.042)
Education (Base = No education)					
Some primary	0.859** (0.336)	0.005 (0.008)	0.797** (0.334)	-0.382 (0.512)	0.809** (0.332)
Primary completed	1.107*** (0.336)	0.015 (0.010)	1.036*** (0.334)	0.328 (0.471)	1.070*** (0.331)
Some secondary	1.013*** (0.365)	0.009 (0.013)	0.947*** (0.362)	0.269 (0.504)	0.928*** (0.355)
Secondary completed	1.425*** (0.343)	0.046*** (0.012)	1.336*** (0.342)	0.806* (0.478)	1.295*** (0.335)
Technical/university	1.763*** (0.346)	0.087*** (0.016)	1.627*** (0.346)	1.185** (0.483)	1.694*** (0.336)
Province					
North Eastern/Coast	0.476** (0.213)	0.030* (0.017)	0.562*** (0.216)	0.488** (0.214)	0.527*** (0.199)
Eastern	0.864*** (0.199)	0.064*** (0.017)	0.928*** (0.211)	0.878*** (0.200)	0.821*** (0.193)
Central	1.353*** (0.216)	0.134*** (0.028)	1.385*** (0.227)	1.374*** (0.216)	1.541*** (0.216)
Rift Valley	-0.686*** (0.222)	-0.020 (0.014)	-0.582** (0.228)	-0.687*** (0.223)	-0.744*** (0.216)
Western/Nyanza	0.209 (0.205)	0.018 (0.015)	0.368* (0.214)	0.214 (0.206)	0.172 (0.197)
Added variables					
Married	Not included	Not included	-0.133 (0.128)	Not included	Not included
Household Size	Not included	Not included	-0.061** (0.029)	Not included	Not included
Urban	Not included	Not included	0.173 (0.137)	Not included	Not included
Household Head	Not included	Not included	0.065 (0.144)	Not included	Not included
Interaction					
Female * Education (2nd Difference)	Not included	Not included			Not included
No education vs. Primary				-0.042* (0.025)	
No education vs. Some sec. education				-0.037 (0.027)	

TABLE A3 (Continued)

Variables	Main model	(1) LPM	(2) Added variables	(3) Interaction	(4) All ages + Income imputed
No education vs. Technical/university				-0.005 (0.029)	
Secondary compl. vs Technical/university				0.044* (0.025)	
Constant	-5.112*** (0.667)	0.043 (0.034)	-5.111*** (0.676)	-4.419*** (0.705)	-5.206*** (0.581)
Observations	7583	7583	7583	7583	8561

Note: $N = 7583$. Robust standard errors are shown in parentheses. Authors' own calculations based on 2018 FinAccess Household Survey.

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

TABLE A4 Gender and formal digital credit usage (propensity score matching)

ATT	Std. Err.	t value	Sample size treatment	Sample size control
-0.014	0.008	1.795	4411	2023

Note: Authors' own calculations based on the FinAccess household survey 2018 using nearest neighbour matching.

ATT = Average treatment effect of treated with female being the treatment. Dependent variable = formal digital credit usage. Number of Blocks = 6; balancing property is satisfied; region of common support is [0.221; 0.849]; standard errors are simulated with 1000 bootstrap replications.

TABLE A5 Gender and formal digital credit usage (Blinder-Oaxaca decomposition)

	Coefficient
Men	0.132*** (0.008)
Women	0.081*** (0.006)
Difference	0.051*** (0.008)
Explained	0.031*** (0.003)
Unexplained	0.020** (0.008)

Note: $N = 7583$. Robust standard errors are shown in parentheses. Included covariates are equal to Model 2. Authors' own calculations based on 2018 FinAccess Household Survey.

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.