BOFIT Discussion Papers
9 • 2020

Francis Osei-Tutu and Laurent Weill

Sex, language, and financial inclusion

BOFIT
THE BANK OF FINLAND
INSTITUTE FOR ECONOMIES
IN TRANSITION

BOFIT Discussion Papers
Editor-in-Chief Zuzana Fungáčová

BOFIT Discussion Papers 9/2020
1.4.2020

Francis Osei-Tutu and Laurent Weill: Sex, language, and financial inclusion

ISBN 978-952-323-321-8, online
ISSN 1456-5889, online

Suomen Pankki
Helsinki 2020

Contents

Francis Osei-Tutu and Laurent Weill

Sex, language, and financial inclusion

Abstract

Reference to gender in language can lead individuals to draw distinctions between genders and reinforce traditional views of gender roles. To test our hypothesis that language gender-marking exerts an influence on the gender gap in financial inclusion, we draw on data for 117 countries in the World Bank's Global Findex database and perform logit estimations at the individual level. We find the gender gap in the probability of owning a formal account, having access to a formal credit, as well as having savings in a formal financial institution is higher for countries with gendered languages than for countries with genderless languages. These findings are confirmed in robustness checks that control for alternative measures of culture and estimations at the country level.

JEL Codes: G21, Z13.
Keywords: financial inclusion, gender, language.

Francis Osei-Tutu, orcid.org/0000-0003-2176-3456. Institut d'Etudes Politiques, Université de Strasbourg, 47 avenue de la Forêt Noire, 67082 Strasbourg Cedex. Email: oseitutu@unistra.fr.

Laurent Weill, orcid.org/0000-0002-8630-1351. Corresponding author. Institut d'Etudes Politiques, Université de Strasbourg, 47 avenue de la Forêt Noire, 67082 Strasbourg Cedex, France. Email: laurent.weill@unistra.fr

Acknowledgments
For valuable comments and suggestions we thank Jean-Gabriel Cousin, Zuzana Fungáčová, Risto Herrala, Gaël Imad'Eddine,Michael Juselius, Frédéric Lobez, Chang Ma, Mikko Mäkinen, Estefania Santacreu-Vasut, Jean-Christophe Statnik as well as participants of the seminars in Bank of Finland (BOFIT) in Helsinki and University of Lille.

1 Introduction

In recent years, the promotion of financial inclusion, measured as the access to and the use of financial services, has taken a prominent place on the agendas of many governments and international organizations. The World Bank, for example, has set the aspirational objective of achieving universal access to a transaction account by 2020. These efforts have been motivated by the fact that financial inclusion is recognized as an important driver of economic development. It not only provides individuals a safe place to save for the future, launch a business or invest in education, but also helps society at large tackle the challenges of reducing poverty and improving health (Dupas and Robinson, 2013; Bruhn and Love, 2014).

A major issue in this debate is the gender gap – specifically, the fact that women continue to have poorer access to financial services than men. The latest wave of Global Findex data for 2017 shows, for example, that 72% of men and 65% of women had bank accounts – a seven-point gender gap that has not changed since the first wave of Global Findex data in 2011 (Demirgüc-Kunt et al., 2018). The gender gap in financial inclusion is an obstacle for women empowerment as it diminishes the economic role of women and their ability to contribute to family support (Hashemi, Schuler and Riley, 1996; Pitt, Khandker and Cartwright, 2006; Swamy, 2014). Thus, understanding the gender gap in financial inclusion is crucial in promoting gender equality.

Despite wide documentation of the gender gap in financial inclusion, the underlying reasons for this gap remain scarcely investigated. Demirgüç-Kunt, Klapper, and Singer (2013) argue that legal discrimination against women (e.g., restrictions in the ability to work or head a household) and gender norms may explain the gender gap. Beck, Behr, and Madestam (2018) provide a behavioral reason, showing that male loan officers charge higher interest rates and grant lower loan amounts to female borrowers, and impliedly that women have lower access to credit in countries with higher shares of male loan officers. Ghosh and Vinod (2017) show the influence of political-, wage-, and education-related determinants for India.

This paper provides a novel explanation for the gender gap in financial inclusion: gender-marking in language. We investigate how language gender-marking influences gender inequalities in the access and use of financial services. Our hypothesis is that languages that require reference to gender lead individuals to draw subtle distinctions between genders. This aspect of language reinforces traditional views of gender roles in the minds of speakers, thereby affecting women's financial inclusion. This hypothesis has its roots in recent research showing that grammatical gender shapes the way people think along gender lines (Boroditsky, 2009).

Languages differ in the use of gender-specific nouns and pronouns. They can utilize a noun assignment based on gender with the masculine or feminine categorization of nouns as in French

("le/la"), which does not exist in English. They can use gender-specific pronouns like "he/she" for the third person pronoun in English. Other languages (e.g. Finnish, Swahili, Mandarin) do not systematically mark gender distinctions. There is no noun assignment or pronoun distinctions based on gender. This linguistic feature can influence inequalities in gender outcomes, i.e. languages that constantly call attention to gender distinctions by discriminating between feminine and masculine nouns and pronouns can influence the perception of distinctions between women and men, thereby influencing societal inequalities.

A few studies have found evidence for this view in economics with the findings that women in countries with a gendered language have lower participation rates in the labor market (Mavisakalyan, 2015), or lower participation on boards of directors and in senior management (Santacreu-Vasut, Shenkar, and Shoham, 2014). Drori et al. (2018) consider whether language gender-marking shapes microfinance outreach to women. They find that gender-marking in language reduces the ratio of female to male borrowers in microfinance organizations.

We test the hypothesis that grammatical gender shapes women's financial inclusion on a sample of roughly 350,000 individuals from 117 countries. Combining individual-level survey data from the Global Findex database for financial inclusion with measures for gender intensity of languages from the World Atlas of Language Structures, we examine whether the gender gap in financial inclusion is greater in countries with highly gendered languages. Three financial inclusion indicators are considered:

- formal ownership of a bank account,
- formal access to bank credit,
- formal saving on a bank account.

We also investigate whether the influence of gender-marking in language affects all types of motives for loans and access to alternative sources of borrowing. This helps in identifying whether language gender-marking constitutes a major obstacle for access to credit.

Our results indicate that language gender-marking affects women's financial inclusion. The gender gap in the probability for having a formal account, access to formal credit, and formal saving is significantly higher in countries with gendered language than in countries with genderless languages. We further find that language gender-marking enhances the gender gap in access to credit for all loan motives and for all sources of borrowing (formal and informal). Our key conclusion is that gendered languages help foster the gender gap in financial inclusion.

Our paper contributes to the literature in two ways. First, we extend the literature on the gender gap in financial inclusion by investigating the influence of language gender-marking. Recent works demonstrate the cultural determinants of gender inequality in economics. Alesina, Giuliano,

and Nunn (2013) show that gender norms are influenced by traditional agricultural practices, specifically plough cultivation. In a related vein, Hansen et al. (2015) argue that higher gender inequality and lower female labor participation are caused by high patriarchal values and beliefs regarding the proper role of women in societies with long agrarian histories.

Second, we add to the discussion on the impact of linguistic structures on economic behavior. This line of research tests the "Sapir-Whorf hypothesis," which states that actions are influenced by language, in economics. This hypothesis of economic behavior influence has been bolstered by evidence from future-tense-marking (Chen, 2013; Mavisakalyan, Tarverdi, and Weber, 2018) and language gender-marking (Santacreu-Vasut, Shenkar, and Shoham, 2014; Mavisakalyan, 2015). To our best knowledge, the influence of language on the gender gap in financial inclusion has never been studied. While Drori et al. (2018) present results for the ability of women to obtain loans in microfinance institutions, we extend the analysis to all dimensions of financial inclusion, including the ownership of a bank account and consider access and use of services in formal financial institutions.

The rest of this paper is organized as follows. Section 2 presents the background of the research question. Section 3 outlines the data and methodology. Section 4 provides estimations results. Section 5 reports the robustness checks. Section 6 concludes.

2 Background

The Sapir-Whorf hypothesis, or *Linguistic Relativity Hypothesis*, states that the language we speak influences our world view (Whorf, 1956). Language, beyond being an important tool for communication, indirectly influences behavior at the subconscious level. While the strong version of this hypothesis, which assumes that language determines thought and controls cognitive processes, has been widely panned, its weak version, which says language constrains cognition, is widely supported (Boroditsky, 2001; Slobin, 2003).

Gender features are encoded into many of the world's nearly 7,000 living languages (Boroditsky, 2011). Linguistic gender differs from biological sex distinctions in that "genders are classes of nouns reflected in the behavior of associated words" (Hockett, 1958, p. 231). Grammatical gender is merely a way of categorizing nouns. For example, the Spanish word for book with definite article, *el libro* (the book), has a masculine grammatical gender without any linkage to the male sex.

The large differences in how languages differ in the way they encode gender can be illustrated by the fact that languages such as Finnish make no distinction as to grammatical gender, while other languages such as Arabic express gender even in the first- or second-person pronoun. English

takes the middle ground; nouns have no grammatical gender but it distinguishes between pronouns in the third-person singular: *he* (masculine) and *she* (feminine).

In gendered languages, the need to constantly consider and mark gender has been found to influence speakers' cognition and how they view objects as having male and female traits. Boroditsky (2009) asked German and Spanish speakers to describe two objects which have opposite gender assignments in the two languages: a *key* which is masculine in German and feminine in Spanish, and a *bridge* which is feminine in German and masculine in Spanish. German speakers described the key using words such as *heavy, hard, metal, serrated,* and *jagged,* whereas Spanish speakers were more likely to say *tiny, intricate, golden, little, lovely,* and *shiny.* In describing the bridge, German speakers used words like *pretty, elegant, peaceful,* and *fragile*, while Spanish speakers said *long, strong, sturdy,* and *big.* This work builds on considerable experimental data. For example, Boroditsky, Schmidt, and Phillips (2003) and Phillips and Boroditsky (2003) tend to show that "even small flukes of grammar, like the seemingly arbitrary assignment of gender to a noun, can have an effect on people's ideas of concrete objects in the world" (Boroditsky, 2009).

In line with this view, scholars and feminists have long argued that gendered languages enforce sex distinctions on its speakers and contribute to sexist outcomes (Stahlberg et al., 2007; Wasserman and Weseley, 2009). The feminist Dale Spender (1985) notes that "there is sexism in language, it does enhance the position of males, and males have had control over the production of cultural forms."

One widespread form of asymmetry is the use of masculine generics, i.e. the use of supposedly gender-neutral masculine pronouns and nouns like "*his*" and "*he*" in statements, even when the referent is not necessarily of a male gender. In this way, maleness is equated to humanness and portrayed as the norm. Using masculine generics has been shown to evoke perceptions and mental representations of men rather than women, and in consequence, put women at a disadvantage (Hamilton et al., 1992; Stahlberg et al., 2001).

A growing body of research in economics confirms the feminist language critique and reveals that gendered languages may highlight the salience of gender distinctions in the minds of speakers, leading to more pronounced inequalities in gender outcomes. Mavisakaylan (2013) examines whether sex-based grammatical gender in languages affect women's labor outcomes in over 100 countries. They find that in countries with a gender-intensive language, women have lower labor force participation. Similarly, Santacreu-Vasut, Shenkar, and Shoham (2013) show that countries with gendered languages are more likely to have political gender quotas and less microfinance support to women (Drori et al., 2018). Van der Velde, Tyrowicz, and Siwinska (2015) find that countries with gender-intensive languages tend to have higher estimates of gender wage gap. At the individual level, Davis and Reynolds (2018) observe that women speaking gendered languages have

lower educational attainment and secondary school completion rate relative to men, while Hicks, Santacreu-Vasut and Shoham (2015) show that households with individuals whose native language marks grammatical gender are more likely to allocate household tasks on the basis of sex. These findings support the view that grammatical gender in languages make a difference in how speakers organize their beliefs about gender and possibly play a role in affecting women's outcomes.

Following this line of reasoning, we question whether the pervasiveness of gender in language influences the ability of men and women to access and use of financial services. In line with the view that language influences actions, we argue that individuals speaking a highly gendered language are more likely to accept prevailing cultural norms and traditional gender roles.

Indeed, for individuals speaking a gendered language, sex-based distinctions are salient in every thought and speech. As explained by Johansson (2005), the origins of gender distinctions in languages are based on evolutionary pressures relating to reproduction, the division of labor, and specialization.[1] Thus, grammatical gender systems may have become embedded in language, and therefore the use of sex-based distinctions in language today may be reflecting historically determined gender-related cultural roles and values.[2]

Further, the direct impact of grammatical gender on cognition (Boroditsky et al., 2003) may reinforce the formation of beliefs and preferences, and thereby influence the behavior of speakers. This may lead women speaking gendered languages to engage less in roles traditionally considered the purview of men such as labor force participation (Mavisakaylan, 2013) and financial responsibilities.

Within the household, as shown by Hicks, Santacreu-Vasut, and Shoham (2015), grammatical gender-marking in language is a strong predictor of the tasks that are engaged in by women and men. We therefore expect that women whose language marks grammatical gender would have a lower probability of owning a formal account, having access to formal credit, and saving in a formal financial institution. This hypothesis accords with the role of language gender-marking in influencing perceptions of gender roles and contributing to gender inequalities.

Guided by our theoretical framework, we propose our testable hypothesis: "Female-male distinctions in language increase the gap in financial inclusion between men and women."

[1] Linguistic features, e.g. grammatical gender marking, have remained stable for a very long time (Wichmann and Holman, 2009), and therefore expected to be exogenous to economic outcomes (Tabellini, 2008).
[2] Alesina, Giuliano, and Nunn (2013) investigate the historical origins of current differences in gender roles in pre-industrial agricultural societies. They conclude that ancestral gender norms still matter today.

3 Data and methodology

3.1 Measuring financial inclusion

We use data on financial inclusion from the World Bank's Global Findex database.[3] This database is obtained from nationally representative surveys of individuals in several countries by the Gallup Inc., in association with its annual Gallup World Poll. Targeting the entire civilian, non-institutionalized population aged 15 and above, the survey randomly selects roughly 1,000 respondents in each economy. Questions are provided in over 140 languages. Some countries have more than 1,000 respondents in a particular poll. The Global Findex contains three waves of data (2011, 2014 and 2017). By employing all three waves, we obtain a fairly broad sample size.[4]

In line with previous studies (e.g. Hannig and Jansen, 2010; Demirgüç-Kunt and Klapper, 2013; Fungacova and Weill, 2015), we focus on the access and use of financial services, including owning an account, borrowing from a financial institution, and saving at a financial institution.[5] We measure financial inclusion by examining three main variables.

Formal account: Does the individual have an account either at a financial institution or through a mobile money provider? Having a *formal account* is the most basic form of financial inclusion as it sets the tone for the use of a diverse range of financial services.

Formal credit: Has the individual borrowed from a financial institution at any time during the past 12 months? The ability to access formal credit is an essential element of financial inclusion especially for households and small businesses who often lack the capital to expand their business activities.

Formal saving: Has the individual saved money in an account with a financial institution at any time during the past 12 months?

In addition to these variables, we also consider several barriers to financial inclusion, the borrower's motivation for taking a loan, and alternative sources of credit.

We examine some perceived barriers preventing individuals from having formal accounts. Respondents are asked to answer the question: "*Please tell me whether each of the following is a reason why you, personally, do not have an account at a bank or another type of formal financial institution."* The survey includes several responses and allows multiple answers. We focus on the barrier: "*because family member already has one*" and assign a dummy equal to one if the respondent answers "yes" to this question, and zero otherwise.

[3] The database is available at the World Bank website: https://datatopics.worldbank.org/financialinclusion.
[4] It is highly unlikely for one individual to participate in more than one of the three surveys, since the data from each wave are processed independently.
[5] While we focus on the access and use of financial services by individuals, Hainz and Nabokin (2019) provide an analysis of access to credit and its determinants for firms.

Next, we analyze the borrower's motivations for taking a loan. Individuals are asked the following question: "*In the past 12 months, have you, by yourself or together with someone else borrowed money for any of the following reasons*?" The responses to this question are: *for education or school fees*, *for medical purposes*, and *for farm or business*. This question relates to either formal or informal credit. To capture motivations for taking only formal credit, the following question is asked: *"Do you, by yourself or together with someone else, currently have a loan you took out from a bank or another type of formal financial institution to purchase a home, an apartment, or land?"*

In addition to the use of formal credit, we delve deeper to explore information on the alternative sources of borrowing. Respondents are asked the question: "*In the past 12 months, have you, by yourself or together with someone else, borrowed any money from any of the following sources?"* We focus on the following responses: *borrowed money from a store*, *borrowed money from family or friends*, and *borrowed money from another private lender*. We then compute the variable *Informal Credit* to measure if a respondent has borrowed any money from any of these three sources in the past 12 months.

3.2 Measuring gender marking in languages

Data on the measures of gender intensity in language come from the World Atlas of Language Structures (WALS: Dryer and Haspelmath, 2013).[6] Following Santacreu-Vasut, Shenkar and Shoham (2014), we employ the four grammatical structure variables related to gender in this database to measure gender intensity of a country's dominant language.[7,8] Each variable captures a different aspect of gender intensity in a language.

The first variable *Sex-Based Index* takes into account whether a gender system is linked to biological sex through a female-male distinction. A language gender system may be classified as sex-based (feminine-masculine distinctions) or non-sex-based (Corbett 2013b). Non-sex-based gender systems apply to languages where gender is not based on natural gender, but rather some notion of animacy. For example, in the Fula language, a member of the Niger-Congo language family,

[6] The recent dataset from Jakiela and Ozier (2019) covers a larger dataset of gendered grammar in the world's languages than the WALS dataset. Nevertheless, we prefer the WALS dataset for two reasons. First, the WALS dataset has a key advantage for our investigation: it contains refined linguistic information that enables us to measure four different aspects of gendered grammar in languages. Second, it is not an issue that the WALS dataset only covers a fraction of the world's languages. We only consider here the dominant language of each country (not all spoken languages of the world) because the Global Findex database only provides information on the country of the respondent. The information on the 117 countries available in the Global Findex database provides a vastly-more-than-adequate sample of respondents for our individual-level estimations.

[7] A language with the highest number of speakers in a country is defined as the dominant language following Encyclopedia Britannica, (2010).

[8] We have done a robustness check, in which we only keep countries where the dominant language is spoken by at least 80% of the population. We obtain similar findings, showing that our results are not influenced by countries in which multiple languages are spoken. These estimations are available upon request.

nouns referring to human females and human males are merged into a common gender. Thus, the gender system in this language is based on the distinction between human and nonhuman. Other non-sex-based gender languages like Swedish and Zulu make distinctions based on animate and inanimate objects. We code the dummy variable *Sex-Based Index* equal to one if a language has a sex-based gender system (e.g. English, Spanish) and zero otherwise (e.g. Swedish, Danish).

The second variable *Number of Genders Index* considers the number of genders a language has. Some languages have multiple genders (Nigerian Fula has around twenty genders), while others like Mandarin or Cantonese have no genders (Corbett, 2013a). Normally, languages that have just two genders, e.g. Arabic and French, possess a "masculine" and "feminine." Languages with three genders, e.g. German and English, include a neuter gender (feminine, masculine, and neuter), and there is no biological sex-related distinction in languages with several genders. We code the dummy variable *Number of Genders Index* equal to one if a language has exactly two genders (e.g. French and Arabic), and zero otherwise (e.g. English and Mandarin).

The third variable, *Gender Pronoun Index,* captures whether a language distinguishes gender in pronouns. Some languages distinguish pronouns along gender lines, meaning they use female/male pronouns when referring to the feminine or masculine, respectively. These distinctions can be made for the third-person pronouns, as well as for the first- and second-person pronouns. For example, English distinguishes between genders only in third-person pronouns (he/she/it), whereas Finnish has no gender distinction in pronouns. In general, languages that have gender distinctions in the first-person pronouns have gender distinctions in the second- or third-person pronouns, or both (Greenberg, 1963). Hence, we compute this variable as a dummy variable which equals one for languages which distinguish gender in the third-person pronouns and also in the first- and/or second-person pronouns (e.g. Spanish and Arabic), and zero otherwise (e.g. English and Finnish).

The fourth variable *Gender Assignment Index* considers how gender is assigned to nouns. In languages, the assignment of genders may depend on the meaning (semantic) of the noun or on its form. In the semantic assignment system, "the meaning of a noun is sufficient to determine its gender, for all or almost all nouns" Corbett (2013c). For example, in Kannada, a language spoken predominantly by people in southern India, nouns that refer to male humans are assigned to the masculine gender, those that refer to female humans are assigned feminine, whereas all other remaining nouns are neuter. This is the case in English, which assigns gender only based on the meaning of the noun, hence a word like "chair" is assigned the neuter gender. On the other hand, some languages assign gender based on both semantic and formal rules. In this assignment system, nouns which are not sex-differentiable may not necessarily be assigned to the neuter gender. Such nouns may be assigned to genders based on morphological and/or phonological (e.g. in Spanish, nouns ending in "*a"* are usually feminine) formal assignment rules. We thus introduce this dummy variable

equal to one if a language assigns gender on both semantic and formal grounds (e.g. French and Russian), and zero otherwise (e.g. English and Cantonese).

Finally, we consider the overall intensity of gender in a language by constructing the *Aggregate Gender Intensity* as the sum of all the four grammatical gender dummies described above. *Aggregate Gender Intensity* captures the pervasiveness of gender distinctions in a language with a value of 4 representing highly gendered languages and 0 for gender neutral languages. For instance, French has a value of 3, since grammatical gender depends on sex-based distinctions (*Sex-Based Index* = 1), has two genders (*Number of Genders Index* = 1), distinguishes gender in the third-person pronouns only (*Gender Pronoun Index* = 0), and assigns gender to nouns on both semantic and formal grounds (*Gender Assignment Index* = 1).

We use *Aggregate Gender Intensity* to create the dummy variable *AGIV* so that we can classify languages in two groups. *AGIV* is equal to one if *Aggregate Gender Intensity* has a value of 3 or 4, corresponding to highly gendered languages. *AGIV* is equal to zero if *Aggregate Gender Intensity* has a value of 0, 1, or 2, corresponding to mildly gendered languages. For example, English is a mildly gendered language with a value of 1, while French, with a value of 3, is a highly gendered language.

3.3 Methodology

We start by testing the outlined hypothesis with logit regressions using the following model specification:

$$X_{ik} = \alpha + \beta * \text{Female}_{ik} + \delta * \text{Individual Controls}_i + \rho * \text{Country Controls}_k + \varepsilon_{ik}, \quad (1)$$

where *X* represents the financial inclusion variable, *i* the individual, and *k* the country. *Female* is a dummy variable equal to one if the individual is a woman, and zero otherwise.

We perform estimations by splitting the countries into two groups based on the intensity of grammatical gender marking for each language gender index. We refer to these two groups as "genderless" (labeled as 0 at the top of the column in the tables of estimations) and "gendered" (labeled as 1 at the top of the column in the tables of estimations). This enables us to identify whether the gender gap in financial inclusion is influenced by the gender intensity of the language by testing if the coefficient for *Female* is significantly lower in gendered countries than in genderless countries.

We include individual-level control variables in line with former literature on the determinants of financial inclusion (Demirgüc-Kunt and Klapper, 2013; Fungacova and Weill, 2015). These variables are provided in the Global Findex database for each respondent. *Age* is the number of

years of the individual. We also compute the square of the respondent's age (Age^2) to control for the possibility of any non-linear relation between age and financial inclusion. The respondent's income level is also considered. We introduce four dummies to capture if the respondent is in the first income quintile (*poorest 20%*), second income quintile (*second 20%*), middle income quintile (*third 20%*), fourth income quintile (*fourth 20%*), and consider the richest income quintile dummy as the omitted variable.

We consider three country-specific control variables in the estimations. We take into account the quality of institutions with the indicator *Rule of Law,* capturing the perceptions of the extent to which people have confidence in and abide by the rules of society. Since access to financial services might depend on the level of economic development, we include the logarithm of GDP per capita (*Log(GDP/capita)*). Finally, we control for population size with the logarithm of the total population who are 15 years or older (*log(population)*) because it accounts for market size affecting the supply of financial services. Data on country-specific control variables are taken from the World Development and World Governance Indicators.

The sample used for the estimations consists of 351,319 observations from 117 countries across the world. Table 1 displays the descriptive statistics for all variables we employ in our study.[9]

4 Results

In this section, we present the results for the main financial inclusion indicators, and give our findings for the determinants of barriers to financial inclusion. We complete the analysis with the factors influencing loan-taking motives and the alternative sources of borrowing.

4.1 Determinants of main financial inclusion indicators

We investigate how language gender marking can influence women's financial inclusion. In all estimations, we present the marginal effects to measure both statistical significance and economic significance. We also report the chi-square test for the difference between coefficients to indicate whether the coefficient for females speaking gendered languages is significantly different from the coefficient for females speaking genderless languages. The estimations for each of the three main financial inclusion variables are reported respectively in Tables 2, 3, and 4, for *formal account,*

[9] The indicators of financial inclusion have different numbers of observations. It can be caused by the fact that the sample for *Formal saving* is a subgroup of the individuals having a formal account, and as such is smaller than the one for *Formal account*. It can also result of missing information on the answers provided by the respondents in the database. Some questions on loan-taking motives and alternative sources of credit were also not asked in the three waves of the database.

formal credit, and *formal saving.* We report for each financial inclusion variable the estimations performed with each of the five grammatical gender indices.

Formal account: We observe that *Female* is significantly negative in all estimations. Being a woman significantly reduces the possibility of having a formal account no matter the language, genderless or gendered. However, the chi-test shows that the coefficient of *Female* is significantly lower with gendered languages than with genderless languages for all the five language gender indices. In other words, the gender gap in financial inclusion is higher in countries with gendered languages than in countries with genderless languages. Thus, grammatical gender-marking enhances the gender gap in account ownership.

We can measure the economic significance of the influence of language gender-marking with the marginal effects. If we consider the aggregate index, we observe that the gender gap is 11.1 percentage points higher in a highly gendered language than in a mildly gendered language.

The estimated coefficients of the control variables are as expected. *Age* and *Age²* have significant effects on formal account, with positive and negative signs, respectively. They show a nonlinear relation between age and formal account suggesting that even though old people are more likely to own an account, this relation diminishes at a certain age. This finding is in line with the result of Fungacova and Weill (2015), who explain this nonlinear relation from the point of a "generational effect," whereby old people have a diminished interest in using financial services or financial institutions are not motivated to seek them out as customers. In accordance with Demirgüç-Kunt and Klapper (2013), we observe that a higher level of income increases the possibility of owning a formal bank account. All four income dummies are significantly negative, with the results showing that moving from a lower income quintile to a higher income quintile is associated with a greater probability of account ownership. Regarding the country-level variables, we point out that *Log(GDP/capita)* is positive and significant, indicating that higher levels of economic development is associated with a higher ownership of bank accounts. *Rule of law* has a positive and significant coefficient, in line with the view that better quality of institutions favors financial inclusion. Finally, higher population size increases the likelihood of owning a bank account.

Formal credit: we also find a significantly negative coefficient for *Female* in all estimations. This suggests that women have a lower access to credit in all countries, whether or not the language marks grammatical gender.

Again, the chi-test shows that the coefficient of *Female* is lower with highly gendered languages than with mildly gendered languages with all five language gender indices. In terms of economic significance, the gender gap in formal credit is 1.5 percentage points higher in a highly gendered language than in a mildly gendered language, when considering the aggregate index. Again, it is economically significant even if the gender gap driven by the language gender marking is lower

than for formal account. Note that the percentage of individuals in the sample is much lower with a formal credit (12.5%) than with a formal account (60.1%). Thus, we show that access to credit is lower for women speaking gendered languages than women speaking genderless languages.

Formal saving: We observe the same findings as for the other financial inclusion indicators. The coefficient of *Female* is significantly negative in all estimations, showing a lower access to formal saving for women in all types of countries. However, the chi-test confirms that this coefficient is lower in countries with gendered languages than in countries with genderless languages. In terms of economic significance, the gender gap in formal saving is 5.1 percentage points higher in a highly gendered language than in a mildly gendered language.

In summary, our results strongly support the view that language gender-marking affects women's financial inclusion. The gender gap in the probability of having a formal account, access to formal credit, and formal saving is significantly higher in countries with gendered languages than in countries with genderless languages. Thus, the hypothesis that sex-based gender systems in languages reinforce traditional gender roles in the minds of speakers, resulting in lower use of financial services for women, is supported by our results.

4.2 Additional estimations

Global Findex database provides additional information which allows us to explore in greater depth the impact of language gender marking on the gender gap in financial inclusion.

4.2.1 Barriers to financial inclusion

First, we examine the reasons why women might not have a bank account. The dataset provides information on the barriers to financial inclusion. Each respondent can answer whether one of the proposed barriers contributes to restrict her/his access to account ownership. Respondents are asked: "*Please tell me whether each of the following is a reason why you, personally, do not have an account at a bank or another type of formal financial institution."* The survey includes the following answer (multiple answers allowed): "*too far away", "lack of documentation", "too expensive", "lack trust", "lack money", "religious reasons", "because family member already has one", "cannot get one",* and *"no need for financial services".*

We focus on the barrier "*because family member already has one*" since this answer may indicate something about the influence of language gender-marking on cultural norms restricting financial inclusion for women. It enables us to understand the deeper causes that influence the barriers hampering women's financial inclusion. The feminist view about the effect of gender systems

in language is the prediction that it positions women to feel inferior to men. For instance, Spender (1985) argues that the view that females should be listed after males because "the male gender was the worthier gender" tends to enhance male power and supremacy. If this claim is true, we should expect that the fact that a family member already owns an account serves as a greater barrier to account ownership for women speaking gendered languages.

We perform estimations explaining the dummy variable equal to one if the respondent answers yes to the question "*because family member already has one*", and zero otherwise. The results are reported in Table 5.

We observe that *Female* is significantly positive in all estimations. Hence, women are more likely to mention the fact that a family member already has an account as an obstacle to not to having one. Women speaking gendered languages are more likely to cite this barrier. The coefficient for *Female* is always higher for women speaking gendered languages. The chi-test shows that the coefficient is significantly higher in gendered languages for all language gender indices. We conclude that women speaking a gender-intensive language are more likely to mention they do not have a bank account because a family member already has one, supporting the view that language can act to exclude women financially through this channel.

This finding shows that women's financial exclusion is voluntary and confirms the feminist claim about the effect of gendered languages on women's outcomes. This result tends to support the view that cultural norms and the view of the prominent role of men in the financial behavior of countries with gendered languages account for the gender gaps in financial inclusion.

4.2.2 Understanding credit behavior

Financial inclusion is essential because it helps the poor and vulnerable individuals finance their education, improve their homes and become entrepreneurs. In this respect, access to credit is a particularly important aspect of financial inclusion.

First, we can question whether language gender marking affects in a similar way the access to loans whatever the motive. The Global Findex database provides information about loan-taking motives. Four potential motives are proposed: "for education", "for medical purposes", "for farm or business", and "to purchase a home or land". We can therefore study whether language gender-marking affects all loan-taking motives similarly.

We redo the estimations by using each of the four loan-taking motivations as the dependent variable. We only perform regressions using *AGIV* as the gender index variable so that we can compare highly gendered and mildly gendered languages. The results are reported in Table 6.

The coefficient for *Female* is significantly negative with highly gendered languages, but not with mildly gendered languages when considering loan-taking motives for education and for medical purposes. It means that women are less likely to obtain a loan than men for both these motives when they speak highly gendered languages, while there is no significant gender gap when they speak mildly gendered languages.

Female is significantly negative for both forms of languages for the loan-taking motives "for farm or business" and "to purchase a home or land". The chi-test, however, shows that the coefficient of *Female* is significantly lower with highly gendered languages than with genderless languages. Our finding suggests that for all four loan-taking motives, the gender gap in the probability of obtaining a loan is higher in countries with highly gendered languages than countries with mildly gendered languages. Thus, language gender-marking affects women's access to credit for all loan-taking motives.

Second, a natural question emerges as to whether the same findings on the influence of language gender-marking on the gender gap in financial inclusion stand when we consider informal borrowing. Up to this point, we have only considered formal borrowing, i.e. loans acquired through formal banking institutions. If our hypothesis is correct, the influence of language should persist no matter the source of borrowing. All forms of access to credit, formal or informal, should be influenced the same way by language gender-marking.

We can investigate this question since we have information on alternative sources of borrowing other than formal credit. We know that if respondents to Global Findex surveys have obtained a loan from "another private lender", "a store", or "family and friends". We then perform estimations by considering each of these alternative sources of borrowing as the dependent variable, and also by considering them all together with the variable "informal credit". We consider only *AGIV* as the gender index variable to focus on the comparison between highly gendered and mildly gendered languages. The estimations are displayed in Table 7.

The results support the view that language gender marking affects all sources of borrowing. For *informal credit* and for loans obtained *from store* and from *family and friends*, the coefficient for *Female* is significantly negative in all estimations. However, the Chi-test shows that the coefficient for *Female* is significantly lower for women speaking highly gendered languages. For loans obtained from *another private lender*, the coefficient for *Female* is significantly negative only for women speaking highly gendered languages.

Our results suggest that women speaking a highly gendered language have a lower probability of obtaining an informal loan than women speaking mildly gendered languages. This finding supports our main conclusion: language gender-marking enhances the gender gap in access to credit. This conclusion stands for formal and informal credit.

5 Robustness checks

This section presents a battery of robustness tests, wherein we include additional measures for culture, perform estimations at the country level, include country fixed effects, and use an alternative definition for the AGIV variable.

5.1 Additional measures for culture

Language is only one aspect of culture. Culture can be broadly defined as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation" (Guiso, Sapienza and Zingales, 2006). Therefore, culture also includes religion, trust, and many other values. If linguistic structures influence economic outcomes, then other cultural dimensions may have similar impact. It is consequently possible that the gender gap in financial inclusion is driven by the estimated effect of language gender systems capturing other cultural aspects. We thus seek to rule out this possibility by performing additional estimations in which we control for alternative culture measures. These additional estimations are reported in Tables 8–10 with alternatively *Formal Account, Formal Credit*, and *Formal Saving* as the dependent variable. In all estimations, we only consider the aggregate gender index with *AGIV* to measure gender intensity.

Religion: Religion has been identified as a key component of culture that shapes the norms of societies (Iyer, 2016). Several studies have shown that religion can influence financial inclusion, notably through the religious prescriptions of Islam for finance (Mohieldin et al., 2011; Demirgüc-Kunt, Klapper and Randall, 2013). We control for religion by including two religion measures at the country level: *Catholic* and *Muslim*. Both these variables are dummy variables equal to one if more than 50% of the inhabitants in a country are respectively Catholics, and Muslims. Data come from the CIA World Factbook. The religion variables are added in the estimations in the first and second columns of each table. We still observe that the coefficient for *Female* is significantly lower when women speak highly gendered languages than when they speak mildly gendered languages. Thus, religion does not drive our results. Interestingly, we observe that both religion variables are significantly negative in most estimations, supporting the view that Catholic and Muslim countries are associated with lower financial inclusion, corroborating the finding for Muslim countries from Demirgüc-Kunt, Klapper and Randall (2013).

Hofstede culture dimensions: The six-dimension terminology of Hofstede (1980, 2001) to characterize culture has been widely adopted to assess the influence of cultural dimensions. We focus on three of these: *Power distance,* which measures the extent to which individuals accept

inequality; *Individualism*, which measures the degree of interdependence a society maintains among its members; and *Masculinity*, which measures the extent to which social gender roles in a society are distinct. In the third and fourth columns of each table, we add the three Hofstede culture variables in our estimations. We still find a significantly negative coefficient for *Female* in all our estimations showing that other aspects of national culture do not drive our results.

Trust and corruption: These two variables have been shown to influence a wide set of economic outcomes including financial development (De Koker and Jentzsch, 2013; Farooq et al., 2013). We therefore take into account trust and corruption in the estimations. *Trust* is assessed with the trust index proposed by La Porta et al. (1997). *Corruption* is measured with Transparency International's corruption perception index, in which higher values indicate less corruption. We add trust and corruption in the estimations performed in the fifth and sixth columns of each table. We again find that the coefficient for *Female* is significantly lower for women speaking highly gendered languages than women speaking mildly gendered languages. Hence, the impact of language gender-marking on the gender gap in financial inclusion is still observed when trust and corruption are taken into account, supporting the robustness of this finding.

Historical use of the plough: Alesina, Giuliano, and Nunn (2013) investigate the historical origins of current differences in gender roles. They argue that the evolution and persistence of gender norms have been influenced by traditional agricultural practices, particularly plough cultivation. Plough cultivation requires significant body and grip strength needed to either pull the plough or control the animal that pulls it, and hence, men tend to have an advantage in farming relative to women. Consistent to this argument, they find evidence that societies that traditionally practiced plough agriculture have lower rates of female labor force participation and a higher prevalence of attitudes favoring gender inequality today.

We can therefore question whether our conclusion for language gender-marking is driven by traditional agricultural practices. To investigate this question, we perform estimations in which we include the plough measure constructed by Alesina, Giuliano and Nunn (*Plough use*). It is the proportion of citizens with ancestors who traditionally used the plough in pre-industrial agriculture. This variable is added in the seventh and eighth columns of all tables. Our key result is preserved: the coefficient for *Female* is significantly lower for women speaking highly gendered languages than women speaking mildly gendered languages. It means that traditional agricultural practices do not drive our results.

Arabic speakers excluded: Arabic is a highly gendered language with an Aggregate Gender Intensity of 4. It represents a substantial share of the languages spoken in our sample (12.84%), and thus could drive our results. The MENA region, which includes a large share of Arabic-speaking individuals, is one of the regions in the world with the largest gender gaps in financial inclusion

(Demirgüc-Kunt, Klapper and Singer, 2013). We perform estimations in the ninth and tenth columns of all tables by excluding Arabic speakers. We still find that the coefficient for *Female* is significantly lower for women speaking highly gendered languages than women speaking mildly gendered languages.

5.2 Country-level estimations

Up to this point, we have performed estimations at the individual level as we link the gender of the individual with her/his level of financial inclusion. We then compare the observed results for women speaking highly gendered languages and women speaking mildly gendered languages.

We can nonetheless check whether our results stand when we perform country-level estimations. A large set of studies on financial inclusion have considered this level for the estimations (e.g. Demirgüc-Kunt and Klapper, 2013) by explaining the share of adults financially included. To this end, we redo our estimations at the country level.

The dependent variables are the percentage share of women (men) reported having a formal account, a formal credit, and a formal saving. This information is directly provided in the Global Findex database. We test alternatively the influence of each of the five gender indices on the aggregate financial inclusion measures. We include the three country-specific control variables formerly used in the individual-level estimations: *Log(GDP/capita), Rule of Law, Log(Population).*

The hypothesis of a gender gap in financial inclusion driven by language gender-marking would be supported if we find that the coefficient of the gender index is more negative when explaining the percentage share of women financially included than when explaining the percentage share of men financially included. It would mean that the gender gap in financial inclusion is higher in countries with gendered languages.

The estimations are reported in Table 11. Our results support our hypothesis. We find that the coefficient of the gender index is significantly negative in all estimations when explaining the percentage shares of women and men financially included, with two exceptions with non-significant coefficients for the percentage share of men financially included. Furthermore, the chi-test shows that the coefficient of the gender index is lower for the percentage share of women financially included than for the percentage share of men financially included in all estimations with an exception with no significant difference. Thus, the country-level estimations corroborate our major conclusion that gendered languages enhance the gender gap in financial inclusion.

5.3 Inclusion of country fixed effects

Country fixed effects are not included in the main estimations since our focus is on the impact of a country-level time-invariant variable, i.e. grammatical gender in language. Country differences have been taken into account through three country-level time-varying variables capturing the quality of institutions, the level of economic development, and population size. It is nonetheless of interest to test the influence of country fixed effects in the estimations, since this inclusion allows partially addressing the omitted variable bias.

To this end, we redo the regressions with country fixed effects. These estimations are displayed in Table 12 with alternatively *Formal Account, Formal Credit,* and *Formal Saving* as the dependent variable. In all estimations, we only consider the aggregate gender index with AGIV to measure gender intensity.

Our key findings are preserved. We still observe that the coefficient of *Female* is significantly lower with highly gendered languages than with mildly gendered languages. Thus, we still find that grammatical gender-marking contributes to enhance the gender gap in women's financial inclusion after including country fixed effects, supporting the robustness of our conclusions.

5.4 Alternative definition of AGIV variable

Up to this point, AGIV is a dummy variable equal to one if the aggregate index *Aggregate Gender Intensity* has a value of 3 or 4, and to zero if *Aggregate Gender Intensity* has a value of 0, 1, or 2. As a consequence, it compares highly gendered and mildly gendered languages. We now redefine AGIV by assigning a value of one to any language with *Aggregate Gender Intensity* greater than zero. This way we oppose genderless languages to all other gendered languages. We perform the estimations with alternatively *Formal Account, Formal Credit,* and *Formal Saving* as the dependent variable. The results are reported in Table 13. In all estimations, we still observe that women speaking highly gendered languages are less likely to be financially included than women speaking genderless languages. Thus, our results hold when we take into account a different definition of the AGIV dummy, supporting the robustness of our findings.

6 Conclusion

In this paper, we tested the hypothesis that language gender-marking influences the gender gap in financial inclusion. To this end, we investigated how language gender-marking influences the probability of a woman obtaining a bank account, having a savings account, and accessing credit. We use data from Global Findex database on a large sample of individuals from 117 countries.

Our key finding is that language gender marking affects women's financial inclusion. The gender gap in the probability of having a formal account, formal credit, and formal saving is greater in countries with gendered languages than in countries with genderless languages. Therefore, our results support the hypothesis that sex-based gender systems in languages reinforce traditional gender roles in the minds of speakers, resulting in lower use of financial services for women. This conclusion is robust to the inclusion of alternative culture indicators and to estimations performed at the country level. We also observe that language gender-marking enhances the gender gap in access to credit for all loan motives and all sources of borrowing, formal or informal. This strengthens our finding that language gender-marking generates obstacles to women's access to credit.

This work provides a fresh view of the gender gap in financial inclusion by showing it has cultural roots anchored in language. Two policy implications with different time horizons arise with respect to reducing the gender gap in financial inclusion. Over the short-term, policy reforms that foster women's financial inclusion should focus on countries with gender-intensive languages. Over the long-term, our study supports calls for reforms that make language more gender-neutral.

References

Alesina, A., Giuliano, P., Nunn, N., 2013. On the Origins of Gender Roles: Women and The Plough. Quarterly Journal of Economics, 128, 469–530.

Beck, T., Behr, P., Madestam, A., 2018. Sex and Credit: Do Gender Interactions Matter for Credit Market Outcomes? Journal of Banking and Finance 87, 380–396.

Beck, T., Demirgüç-Kunt, A., Levine, R., 2007. Finance, Inequality and the Poor. Journal of Economic Growth, 12, 27–49.

Boroditsky, L., 2009. How Does Our Language Shape The Way We Think. What's Next 116–129.

Boroditsky, L., 2011. How Language Shapes Thought. Scientific American, 304, 62–65.

Boroditsky, L., 2001. Does Language Shape Thought? Mandarin and English Speakers' Conceptions of Time. Cognitive Psychology, 43, 1–22.

Boroditsky, L., Schmidt, L.A., Phillips, W., 2003. Sex, Syntax, and Semantics. In: Gentner, D., Goldin-Meadow, S. (eds.), Language in Mind: Advances in the Study of Language and Cognition, 61–79. Cambridge, MA: MIT Press.

Bruhn, M., Love, I., 2014. The Real Impact of Improved Access to Finance: Evidence from Mexico. Journal of Finance LXIX, 1347–1376.

Chen, M.K., 2013. The Effect of Language on Economic Behaviour: Evidence from Savings Rates, Health Behaviours, and Retirement Assets. American Economic Review 103, 690–731.

Corbett, G.G., 2013a. Number of Genders. In: Dryer, M.S., Haspelmath, M. (eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology, Chapter 30.

Corbett, G.G., 2013b. Sex-Based and Non-Sex-Based Gender Systems. In: Dryer, M.S., Haspelmath, M. (eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology, Chapter 31.

Corbett, G.G., 2013c. Systems of Gender Assignment. In: Dryer, M.S., Haspelmath, M. (eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology, Chapter 32.

Davis, L., Reynolds, M., 2018. Gendered Language and the Educational Gender Gap. Economics Letters 168, 46–48.

De Koker, L., Jentzsch, N., 2013. Financial Inclusion and Financial Integrity: Aligned Incentives? World Development, 44, 267–280.

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, 279–340.

Demirguc-Kunt, A., Klapper, L., Randall, D., 2013. Islamic Finance and Financial Inclusion: Measuring Use of and Demand for Formal Financial Services Among Muslim Adults. The World Bank.

Demirgüc-Kunt, A., Klapper, L., Singer, D., 2013. Financial Inclusion and Legal Discrimination against Women: Evidence from Developing Countries. World Bank Research Paper 6416.

Demirgüc-Kunt, A., Klapper, L., Singer, D., Ansar, S., Hess, J., 2018. The Global Findex Database 2017: Measuring Financial Inclusion and The Fintech Revolution. The World Bank.

Drori, I., Manos, R., Santacreu-Vasut, E., Shenkar, O., Shoham, A., 2018. Language and Market Inclusivity for Women Entrepreneurship: The Case of Microfinance. Journal of Business Venturing, 33, 395–415.

Dryer, M.S., Haspelmath, M., 2013. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Dupas, P., Robinson, J., 2013. Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. American Economic Journal. Applied Economics 5, 163–92.

Encyclopedia Britannica, 2010. Encyclopedia Britannica, Encyclopedia Britannica Inc., Chicago, IL

Farooq, A., Shahbaz, M., Arouri, M., Teulon, F., 2013. Does Corruption Impede Economic Growth in Pakistan? Economic Modelling 35, 622–633.

Fungacova, Z., Weill, L., 2015. Understanding Financial Inclusion in China. China Economic Review 34, 196–206.

Ghosh, S., Vinod, D., 2017. What Constrains Financial Inclusion for Women? Evidence from Indian Micro Data. World Development 92, 60–81.

Greenberg, J.H., 1963. Some Universals of Grammar with Particular Reference to The Order of Meaningful Elements. Universals of Language 2, 73–113.

Guiso, L., Sapienza, P., Zingales, L., 2006. Does Culture Affect Economic Outcomes? Journal of Economic Perspectives 20, 2, 23–48.

Hainz, C., Nabokin, T., 2019. Access to Credit and Its Determinants: A Comparison of Survey-Based Measures. Economics of Transition and Institutional Change, forthcoming.

Hamilton, M.C., Hunter, B., Stuart-Smith, S., 1992. Jury instructions worded in the masculine generic: Can a woman claim self-defense when "he" is threatened? In: Chrisler, J.C., Howard D. (eds.), New directions in feminist psychology: Practice, theory, and research, 169–178. New York: Springer.

Hannig, A., Jansen, S., 2010. Financial Inclusion and Financial Stability: Current Policy Issues, ADBI Working Paper No. 259.

Hansen, C.W., Jensen, P.S., Skovsgaard, C.V., 2015. Modern Gender Roles and Agricultural History: The Neolithic Inheritance. Journal of Economic Growth 20, 365–404.

Hashemi, S., Schuler, S., Riley, A., 1996. Rural Credit Programs and Women's Empowerment in Bangladesh. World Development 24, 4, 635–653.

Hicks, D.L., Santacreu-Vasut, E., Shoham, A., 2015. Does Mother Tongue Make for Women's Work? Journal of Economic Behavior & Organization 110, 19–44.

Hockett, C.F., 1958. A Course in Modern Linguistics. Oxford and Ibh Publishing Co.; New Delhi.

Hofstede, G., 1980. Culture's Consequences: International Differences in Work-Related Values. Sage. Thousand Oaks.

Hofstede, G., 2001. Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations. Sage, Thousand Oaks.

Iyer, S., 2016. The New Economics of Religion. Journal of Economic Literature. 54, 2, 395–441.

Jakiela, P., Ozier, O., 2019. Gendered Language. CGD Working Paper 500, Washington DC: Center for Global Development.

Johansson, S., 2005. Origins of Language: Constraints on Hypotheses. Amsterdam: John Benjamins.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R.W., 1997. Legal Determinants of External Finance. Journal of Finance 52, 1131–1150.

Mavisakalyan, A., 2015. Gender in Language and Gender in Employment. Oxford Development Studies 43, 403–424.

Mohieldin, M., Iqbal, Z., Rostom, A., Fu, X., 2011. The Role of Islamic Finance in Enhancing Financial Inclusion in Organization of Islamic Cooperation (OIC) Countries. The World Bank.

Phillips, W., Boroditsky, L., 2003. Can Quirks of Grammar Affect the Way You Think? Grammatical Gender and Object Concepts. In: Proceedings of the 25th Annual Meeting of the Cognitive Science Society, 928–933.

Pitt, M.M., Khandker, S.R., Cartwright, J., 2006. Empowering Women with Micro Finance: Evidence from Bangladesh. Economic Development and Cultural Change 54, 791–831.

Santacreu-Vasut, E., Shenkar, O., Shoham, A., 2014. Linguistic Gender Marking and Its International Business Ramifications. Journal of International Business Studies 45, 1170–1178.

Slobin, D.I., 2003. Language and thought online: Cognitive consequences of linguistic relativity. Language in Mind: Advances in the Study of Language and Thought 157, 178.

Spender, D., 1985, Man Made Language, 2nd edition, New York: Routledge.Stahlberg, D., Braun, F., Irmen, L., Sczesny, S., 2007. Representation of The Sexes in Language. Social Communication 163–187.

Stahlberg, D., Sczesny, S., Braun, F., 2001. Name Your Favorite Musician: Effects of Masculine Generics and of Their Alternatives in German. Journal of Language and Social Psychology 20, 464–469.

Swamy, V., 2014. Financial Inclusion, Gender Dimension, and Economic Impact on Poor Households. World Development 56, 1–15

Tabellini, G., 2008. Institutions and Culture. Journal of the European Economic Association 6, 255–294.

Van der Velde, L., Tyrowicz, J., Siwinska, J., 2015. Language and (the estimates of) the Gender Wage Gap. Economics Letters 136, 165–170.

Wasserman, B.D., Weseley, A.J., 2009. ¿Qué? Quoi? Do Languages With Grammatical Gender Promote Sexist Attitudes? Sex Roles 61, 634.

Wichmann, S., Holman, E.W., 2009. Temporal Stability of Linguistic Typological Features. Lincom Europa Munich.

Whorf, B.L., 1956. Language, Thought and Reality. Selected Writings of Benjamín Lee Whorf. Cambridge, MA: MIT Press.

Tables

Table 1 Descriptive statistics

	Observations	Mean	Std. Dev.
Main indicators of financial inclusion			
Formal account	351,319	0.601	0.49
Formal credit	348,948	0.125	0.331
Formal saving	277,357	0.31	0.462
Barrier to financial inclusion			
Family member has an account	144,401	0.189	0.391
Loan-taking motivations			
For education	190,923	0.063	0.242
For medical purposes	311,786	0.101	0.301
For farm or business	235,750	0.059	0.235
To purchase a home or land	349,057	0.131	0.338
Alternative sources of credit			
Another private lender	227,658	0.036	0.187
Family and friends	348,788	0.214	0.41
A store	226,774	0.097	0.296
Informal credit	349,474	0.259	0.438
Language gender variables			
Sex-based index	351,319	0.672	0.47
Number of genders index	351,319	0.446	0.497
Gender pronoun index	333,496	0.327	0.469
Gender assignment index	255,074	0.677	0.468
Aggregate gender index (AGI)	243,189	2.36	1.68
Individual characteristics			
Female	351,319	0.544	0.498
Age	351,319	42.644	17.780
Income – poorest 20%	351,319	0.168	0.374
Income – second 20%	351,319	0.18	0.384
Income – third 20%	351,319	0.195	0.396
Income – fourth 20%	351,319	0.211	0.408
Income – richest 20%	351,319	0.246	0.431
Country variables			
Log (GDP/capita)	351,319	8.921	1.353
Rule of law	351,319	0.057	1.0
Log (population)	351,319	16.501	1.686

This table presents the descriptive statistics of the variables used in the estimations.

Table 2 Formal account

	Sex-based index		Number of genders index		Gender pronoun index		Gender assignment index		AGIV	
	0	1	0	1	0	1	0	1	Mildly gendered	Highly gendered
Female	–0.049***	–0.086***	–0.029***	–0.13***	–0.038***	–0.135***	–0.039***	–0.114***	–0.03***	–0.141***
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age	0.015***	0.023***	0.016***	0.023***	0.017***	0.023***	0.011***	0.022***	0.013***	0.023***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Age^2	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income-poorest 20%	–0.301***	–0.300***	–0.26***	–0.31***	–0.257***	–0.298***	–0.272***	–0.299***	–0.272***	–0.29***
	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)	(0.004)	(0.006)	(0.004)
Income-second 20%	–0.24***	–0.243***	–0.199***	–0.264***	–0.2***	–0.257***	–0.219***	–0.252***	–0.218***	–0.247***
	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)	(0.004)	(0.006)	(0.004)
Income-third 20%	–0.186***	–0.185***	–0.146***	–0.209***	–0.146***	–0.204***	–0.159***	–0.202***	–0.159***	–0.199***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.004)	(0.005)	(0.004)
Income-fourth 20%	–0.113***	–0.114***	–0.087***	–0.132***	–0.091***	–0.125***	–0.096***	–0.128***	–0.099***	–0.126***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.005)	(0.006)	(0.004)	(0.005)	(0.004)
Log (GDP/capita)	0.1***	0.15***	0.105***	0.157***	0.109***	0.184***	0.188***	0.142***	0.136***	0.156***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)
Rule of law	0.268***	0.155***	0.17***	0.175***	0.172***	0.072***	0.054***	0.167***	0.132***	0.138***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)	(0.002)	(0.003)	(0.003)
Log (population)	0.027***	0.017***	0.012***	0.03***	0.013***	–0.014***	0.002	0.03***	0.01***	0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	115,372	235,947	194,548	156,771	224,556	108,940	82,366	172,708	116,213	126,976
Pseudo R^2	0.295	0.269	0.313	0.233	0.314	0.182	0.34	0.219	0.32	0.181
Log likelihood	–55661.31	–114542.45	–85126.04	–83208.49	–98173.59	–61571.31	–34946.15	–93419.3	–50557.78	–71926.41
χ^2	106.92***		443.18***		342.85***		120.98***		399.86***	

This table presents the results of logit estimations examining the relation between language gender systems and women's financial inclusion. The dependent variable is "Formal Account." Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a genderless language and a gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 3 Formal credit

	Sex-based index		Number of genders index		Gender pronoun index		Gender assignment index		AGIV	
	0	1	0	1	0	1	0	1	Mildly gendered	Highly gendered
Female	–0.016***	–0.02***	–0.015***	–0.024***	–0.015***	–0.024***	–0.014***	–0.02***	–0.009***	–0.024***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Age	0.016***	0.014***	0.016***	0.012***	0.015***	0.012***	0.018***	0.011***	0.016***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age^2	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income-poorest 20%	–0.029***	–0.042***	–0.032***	–0.044***	–0.03***	–0.049***	–0.034***	–0.041***	–0.033***	–0.043***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Income-second 20%	–0.019***	–0.03***	–0.02***	–0.034***	–0.019***	–0.04***	–0.026***	–0.034***	–0.026***	–0.034***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Income-third 20%	–0.014***	–0.021***	–0.014***	–0.023***	–0.013***	–0.028***	–0.017***	–0.025***	–0.017***	–0.024***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Income-fourth 20%	–0.006**	–0.013***	–0.008***	–0.013***	–0.007***	–0.016***	–0.011***	–0.014***	–0.013***	–0.013***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Log (GDP/capita)	0.005***	0.021***	0.004***	0.017***	0.008***	0.019***	0.012***	0.016***	–0.001	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
Rule of law	0.014***	0.012***	0.019***	0.014***	0.018***	0.014***	0.007**	0.01***	0.029***	0.02***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)	(0.002)	(0.001)
Log (population)	–0.009***	–0.005***	–0.004***	–0.007***	–0.007***	–0.018***	–0.014***	–0.006***	–0.006***	–0.012
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Observations	114,629	234,319	193,174	155,774	222,872	108,324	81,942	171,607	115,505	126,185
Pseudo R^2	0.043	0.057	0.045	0.059	0.048	0.063	0.048	0.054	0.048	0.061
Log likelihood	–42383.83	–82355.73	–73347.89	–51316.09	–83318.69	–35887.56	–33429.37	–54643.37	–44150.1	–40789.03
χ^2	6.83***		44.58***		30.00***		25.52***		64.21***	

This table presents the results of logit estimations examining the relation between language gender systems and women's financial inclusion. The dependent variable is "Formal Credit". Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a genderless language and a gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 4 Formal saving

	Sex-based index		Number of genders index		Gender pronoun index		Gender assignment index		AGIV	
	0	1	0	1	0	1	0	1	Mildly gendered	Highly gendered
Female	–0.016***	–0.048***	–0.015***	–0.061***	–0.022***	–0.059***	–0.013***	–0.054***	–0.014***	–0.065***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)
Age	0.012***	0.012***	0.013***	0.011***	0.014***	0.01***	0.011***	0.01***	0.012***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Age^2	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***	–0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income-poorest 20%	–0.196***	–0.201***	–0.219***	–0.179***	–0.218***	–0.155***	–0.257***	–0.162***	–0.224***	–0.16***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
Income-second 20%	–0.159***	–0.173***	–0.18***	–0.154***	–0.181***	–0.134***	–0.209***	–0.143***	–0.19***	–0.137***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)	(0.002)	(0.004)	(0.002)
Income-third 20%	–0.114***	–0.132***	–0.129***	–0.12***	–0.129***	–0.108***	–0.141***	–0.113***	–0.134***	–0.108***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.002)	(0.004)	(0.003)
Income-fourth 20%	–0.07***	–0.089***	–0.079***	–0.084***	–0.083***	–0.074***	–0.084***	–0.076***	–0.082***	–0.075***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.004)	(0.003)
Log (GDP/capita)	0.044***	0.073***	0.054***	0.058***	0.064***	0.067***	0.098***	0.051***	0.048***	0.056***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.001)	(0.003)	(0.002)
Rule of law	0.16***	0.12***	0.163***	0.098***	0.159***	0.026***	0.124***	0.085***	0.168***	0.056***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.007)	(0.002)	(0.004)	(0.002)
Log (population)	0.044***	0.012***	0.04***	0.006***	0.031***	–0.005***	0.041***	0.009***	0.044***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	92,434	184,923	156,040	121,317	179,590	83,387	67,212	134,324	94,538	97,396
Pseudo R^2	0.23	0.171	0.207	0.15	0.208	0.103	0.232	0.138	0.226	0.107
Log likelihood	–43510.6	–95397.31	–80208.72	–58814.42	–92262.33	–38676.43	–34263.15	–63941.83	–47375.76	–45952.36
χ^2	54.57***		223.49***		187.68***		136.86***		232.43***	

This table presents the results of logit estimations examining the relation between language gender systems and women's financial inclusion. The dependent variable is "Formal Saving". Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a genderless language and a gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 5 Barrier to financial inclusion

	Sex-based index		Number of genders index		Gender pronoun index		Gender assignment index		AGIV	
	0	1	0	1	0	1	0	1	Mildly gendered	Highly gendered
Female	0.028***	0.043***	0.026***	0.046***	0.033***	0.045***	0.024***	0.04***	0.019***	0.049***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,917	91,484	67,835	76,566	77,749	59,516	29,518	86,846	43,371	67,887
Pseudo R^2	0.091	0.06	0.084	0.061	0.092	0.079	0.097	0.067	0.093	0.064
Log likelihood	–21095.89	–43711.98	–29191.69	–35715.88	–35163.16	–25685.72	–13107.1	–38146.18	–18233.88	–30955.57
χ^2	2.91*		15.73***		14.81***		10.88***		32.99***	

This table presents the results of logit estimations examining the relation between language gender systems and a perceived barrier to women's account ownership. The dependent variable is "Family member already has an account". All controls represent the full set of individual and country level control variables used in Table 2. Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a genderless language and a gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 6 Loan-taking motives

	For education		For medical purposes		For farm or business		To purchase a home or land	
	AGIV		AGIV		AGIV		AGIV	
	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered
Female	0.000 (0.002)	–0.009*** (0.002)	–0.002 (0.002)	–0.013*** (0.002)	–0.016*** (0.001)	–0.03*** (0.002)	–0.008*** (0.002)	–0.025*** (0.001)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	62,646	72,883	102,821	117,250	79,471	84,150	115,517	126,224
Pseudo R^2	0.035	0.017	0.055	0.027	0.068	0.035	0.162	0.111
Log likelihood	–14421.17	–18978.05	–32312.47	–41440.73	–18265.97	–20230.44	–39760.68	–35252.74
χ^2	10.64***		14.22***		25.34***		124.31***	

This table presents the results of logit estimations examining the relation between language gender systems and loan taking motivations for women. The loan taking motivations are: "For education", "For medical purposes", "For farm or business", "To purchase a home or land". Each motivation for accessing loan is a dependent variable and is presented at the top of each column. All controls represent the full set of individual and country level control variables used in Table 3. Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a mildly gendered language and a highly gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 7 Alternative sources of borrowing

	Another private lender		From store		Family and friends		Informal credit	
	AGIV		AGIV		AGIV		AGIV	
	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered
Female	–0.001	–0.019***	–0.008***	–0.023***	–0.017***	–0.051***	–0.02***	–0.057***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	75,300	81,555	75,303	80,703	115,525	126,020	115,679	126,345
Pseudo R^2	0.059	0.02	0.028	0.016	0.067	0.031	–64697.77	0.022
Log likelihood	–10322.6	–16824.63	–22928.96	–25025.21	–57760.93	–64387.86	0.044	–71358.89
χ^2	45.14***		25.81***		99.03***		107.82***	

This table presents the results of logit estimations examining the relation between language gender systems and alternatives sources of borrowing available to women. The alternative borrowing sources are: "Another private lender", "From store", "Family and friends", and "Informal credit." Each alternative source of borrowing is a dependent variable and is presented at the top of each column. All controls represent the full set of individual and country level control variables used in Table 3. Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a mildly gendered language and a highly gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 8 Formal account

	AGIV		AGIV		AGIV		AGIV		AGIV (Excluding Arabic speaking countries)	
	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered
Female	–0.03*** (0.003)	–0.143*** (0.003)	–0.018*** (0.002)	–0.128*** (0.004)	–0.017*** (0.002)	–0.144*** (0.004)	–0.024*** (0.003)	–0.124*** (0.005)	–0.03*** (0.003)	–0.09*** (0.004)
Catholic	–0.103*** (0.005)	–0.377*** (0.005)								
Muslim	0.000 (0.005)	–0.365*** (0.006)								
Power distance			–0.002*** (0.000)	0.000 (0.000)						
Individualism			–0.001*** (0.000)	0.007*** (0.000)						
Masculinity			–0.001*** (0.000)	–0.004*** (0.000)						
Corruption					0.001*** (0.000)	–0.005*** (0.000)				
Trust					0.186*** (0.01)	0.388*** (0.037)				
Plough use							–0.148*** (0.004)	0.125*** (0.006)		
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	113,593	126,976	78,148	93,888	74,282	67,068	76,279	62,924	116,213	81,853
Pseudo R^2	0.332	0.203	0.287	0.183	0.281	0.199	0.273	0.216	0.33	0.172
Log likelihood	–48837.26	–70056.58	–29229.5	–53104.35	–28400.17	–37244.84	–32442.12	–34161.58	–50557.78	–46944.16
χ^2	390.64***		178.65***		237.43***		149.38***		78.77***	

This table presents the results of logit estimations examining the relation between language gender systems and women's financial inclusion. The dependent variable is "Formal Account." All controls represent the full set of individual- and country-level control variables used in Table 3. Definitions of variables are provided in the Appendix. χ2 (chi-squared test) compares the coefficients for Female speaking a mildly gendered language and a highly gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 9 Formal credit

	AGIV		AGIV		AGIV		AGIV		AGIV (Excluding Arabic speaking countries)	
	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered
Female	–0.009*** (0.002)	–0.028*** (0.002)	–0.015*** (0.002)	–0.027*** (0.002)	–0.017*** (0.002)	–0.023*** (0.002)	–0.015*** (0.002)	–0.024*** (0.002)	–0.009*** (0.002)	–0.018*** (0.002)
Catholic	–0.025*** (0.002)	–0.016*** (0.002)								
Muslim	–0.016*** (0.003)	–0.054*** (0.003)								
Power distance			–0.001*** (0.000)	–0.001*** (0.000)						
Individualism			0.001*** (0.000)	–0.000 (0.000)						
Masculinity			–0.002*** (0.000)	0.002*** (0.000)						
Corruption					–0.000*** (0.000)	–0.000*** (0.000)				
Trust					0.053*** (0.01)	0.098*** (0.016)				
Plough use							–0.008*** (0.003)	–0.018*** (0.003)		
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112,887	126,185	77,626	93,287	73,837	66,668	75,797	62,467	115,505	81,303
Pseudo R^2	0.052	0.069	0.066	0.061	0.056	0.074	0.052	0.069	0.048	0.055
Log likelihood	–42703.39	–40433.47	–29531.33	–31490.33	–28949.25	–20665.27	–29375.42	–20257.39	–44150.1	–27368.04
χ^2	100.22***		26.78***		15.82***		18.85***		16.04***	

This table presents the results of logit estimations examining the relation between language gender systems and women's financial inclusion. The dependent variable is "Formal Credit." All controls represent the full set of individual- and country-level control variables used in Table 3. Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a mildly gendered language and a highly gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 10 Formal saving

	AGIV		AGIV		AGIV		AGIV		AGIV (Excluding Arabic speaking countries)	
	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered
Female	–0.014*** (0.003)	–0.066*** (0.003)	–0.01** (0.004)	–0.066*** (0.003)	–0.013*** (0.004)	–0.07*** (0.004)	–0.013*** (0.004)	–0.057*** (0.004)	–0.014*** (0.003)	–0.056*** (0.003)
Catholic	0.001 (0.004)	–0.092*** (0.004)								
Muslim	–0.107*** (0.006)	–0.107*** (0.004)								
Power distance			–0.002*** (0.000)	–0.000 (0.000)						
Individualism			–0.002*** (0.000)	0.003*** (0.000)						
Masculinity			–0.001*** (0.000)	0.001*** (0.000)						
Corruption					–0.002*** (0.000)	–0.001*** (0.000)				
Trust					0.302*** (0.018)	0.398*** (0.029)				
Plough use							–0.132*** (0.007)	0.063*** (0.006)		
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	91,919	97,396	64,565	72,098	63,249	52,572	62,304	47,246	94,538	62,469
Pseudo R^2	0.225	0.113	0.156	0.107	0.204	0.142	0.206	0.155	0.226	0.109
Log likelihood	–46385.77	–45631.31	–37587.46	–35857.57	–34391.23	–24245.12	–33171.2	–21625.6	–47375.76	–30271.31
χ^2	240.74***		180.78***		184.39***		102.03***		116.10***	

This table presents the results of logit estimations examining the relation between language gender systems and women's financial inclusion. The dependent variable is "Formal Saving." All controls represent the full set of individual- and country-level control variables used in Table 3. Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients for Female speaking a mildly gendered language and a highly gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 11 Country-level estimations

	Female share	Male share	Female share	Male share	Female share	Male share	Female share	Male share	Female share	Male share
	Sex-based index		Number of genders index		Gender pronoun index		Gender assignment index		AGIV	
Panel A: Formal credit										
Gender index	–0.072***	–0.04**	–0.126***	–0.07***	–0.164***	–0.096***	–0.102***	–0.054**	–0.047***	–0.03***
	(0.02)	(0.019)	(0.017)	(0.017)	(0.019)	(0.019)	(0.024)	(0.023)	(0.006)	(0.006)
Observations	328	328	328	328	321	321	235	235	223	223
R^2	0.760	0.752	0.786	0.761	0.787	0.758	0.749	0.744	0.766	0.751
χ^2	10.75***		36.79***		39.64***		14.68***		18.84***	
Panel B: Formal credit										
Gender index	–0.025***	–0.027***	–0.028***	–0.015*	–0.023***	–0.007	–0.06***	–0.047***	–0.013***	–0.009***
	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.01)	(0.003)	(0.003)
Observations	328	328	328	328	321	321	235	235	223	223
R^2	0.193	0.280	0.204	0.264	0.212	0.279	0.304	0.346	0.277	0.329
χ^2	0.12		13.81***		16.18***		7.56**		10.78***	
Panel C: Formal saving										
Gender index	–0.057***	–0.032**	–0.052***	–0.018	–0.058***	–0.026**	–0.067***	–0.037**	–0.022***	–0.013***
	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)	(0.013)	(0.014)	(0.015)	(0.004)	(0.004)
Observations	328	328	328	328	321	321	235	235	223	223
R^2	0.741	0.739	0.740	0.736	0.738	0.736	0.735	0.726	0.732	0.720
χ^2	25.08***		47.13***		36.43***		23.53***		36.44***	

This table presents the results of OLS regressions examining the relation between language gender systems and women's financial inclusion. The dependent variables in Panels A, B, and C are respectively "Formal Account", "Formal Credit" and "Formal Saving." Definitions of variables are provided in the Appendix. $\chi 2$ (chi-squared test) compares the coefficients of the Gender Index for the financial inclusion of female and male. Under the null hypothesis, the difference between the two coefficients is zero. Standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Table 12 Robustness check with country fixed effects

	Formal account		Formal credit		Formal saving	
	AGIV		AGIV		AGIV	
	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered	Mildly gendered	Highly gendered
Female	–0.024*** (0.002)	–0.107*** (0.002)	–0.014*** (0.002)	–0.03*** (0.002)	–0.01*** (0.003)	–0.056*** (0.002)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	116,213	126,976	115,505	126,185	94,538	97,396
Pseudo R^2	0.381	0.25	0.097	0.093	0.265	0.165
Log likelihood	–46658.58	–65873.13	–41886.2	–39384.7	–44950.55	–42991.76
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
χ^2	401.93***		70.77***		237.55***	

This table presents the results of logit estimations examining the relation between language gender-marking and women's financial inclusion. The dependent variable is presented at the top of each column. All controls represent the full set of individual- and country-level control variables used in Table 2. Definitions of variables are provided in the Appendix. χ^2 (chi-squared test) compares the coefficients for Female speaking a mildly gendered language and a highly gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 13 Robustness check with alternative definition of AGIV

	Formal account		Formal credit		Formal saving	
	AGIV		AGIV		AGIV	
	Genderless languages	Gendered languages	Genderless languages	Gendered languages	Genderless languages	Gendered languages
Female	–0.066*** (0.005)	–0.103*** (0.003)	–0.012*** (0.003)	–0.019*** (0.001)	–0.017*** (0.004)	-0.051*** (0.002)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	57,901	185,288	57,575	184,115	45,894	146,040
Pseudo R^2	0.243	0.26	0.051	0.063	0.2001	0.177
Log likelihood	–30057.28	–94200.61	–22133.24	–62475.23	–21520.36	–72147.99
χ^2	46.47***		12.06***		41.53***	

This table presents the results of logit estimations examining the relation between language gender-marking and women's financial inclusion. The dependent variable is presented at the top of each column. All controls represent the full set of individual- and country-level control variables used in Table 2. Definitions of variables are provided in the Appendix. The alternative definition of AGIV here is equal to one if Aggregate Gender Intensity is greater than zero (gendered languages), and zero otherwise (genderless languages). χ^2 (chi-squared test) compares the coefficients for Female speaking a genderless language and female speaking a gendered language. Under the null hypothesis, the difference between the two coefficients is zero. Estimated marginal effects are reported and standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Appendix

Variable	Definition and source
Dependent variables	
Formal account	Dummy variable equal to one if individual has an account either at a financial institution or through a mobile money provider; zero otherwise. Source: Global Findex Database.
Formal credit	Dummy variable equal to one if individual borrowed from a financial institution during the past 12 months; zero otherwise. Source: Global Findex Database.
Formal saving	Dummy variable equal to one if individual saved using an account at a financial institution during the past 12 months; zero otherwise. Source: Global Findex Database.
Family member has an account	Dummy variable equal to one if respondent does not have an account at a bank or another type of formal financial institution because a family member already has an account; zero if respondent has some other reason for not having an account. Source: Global Findex Database.
For education	Dummy variable equal to one if respondent, or respondent jointly with someone else, borrowed money for education or school fees; zero otherwise. Source: Global Findex Database.
For medical purposes	Dummy variable equal to one if respondent, or respondent jointly with someone else, borrowed money for medical purposes; zero otherwise. Source: Global Findex Database.
For farm or business	Dummy variable equal to one if respondent, or respondent jointly with someone else, borrowed money for farm or business; zero otherwise. Source: Global Findex Database.
To purchase a home or land	Dummy variable equal to one if respondent took out a loan from a bank or another type of formal financial institution to purchase a home, an apartment, or land; zero otherwise. Source: Global Findex Database.
Another private lender	Dummy variable equal to one if respondent, or respondent jointly with someone else, borrowed money from another private lender; zero otherwise. Source: Global Findex Database.
Family and friends	Dummy variable equal to one if respondent, or respondent jointly with someone else, borrowed money from family and friends; zero otherwise. Source: Global Findex Database.
A store	Dummy variable equal to one if respondent, or respondent jointly with someone else, borrowed money from a store; and zero otherwise. Source: Global Findex Database.
Informal credit	Dummy variable equal to one if respondent, or respondent jointly with someone else, borrowed any money from either another private lender, family and friends, or a store and zero otherwise. Source: Global Findex Database.
Independent variables	
Sex-based index	Dummy variable equal to one if the language has a sex-based gender system, zero otherwise. Source: World Atlas of Language Structures.
Number of genders index	Dummy variable equal to one if the language has exactly two genders, zero otherwise. Source: World Atlas of Language Structures.

Gender pronoun index	Dummy variable equal to one if a language distinguishes gender in the third, first and/or second person pronouns, zero otherwise. Source: World Atlas of Languages Structures.
Gender assignment index	Dummy variable equal to one if a language assigns gender on both semantic and formal grounds, zero otherwise. Source: World Atlas of Languages Structures.
Aggregate gender index	The sum of sex-based index, number of genders index, gender pronoun index and gender assignment index.
Female	Dummy variable equal to one if the individual is a woman and zero otherwise. Source: Global Findex Database.
Control variables	
Age	The number of years of the individual. Source: Global Findex Database.
Income-poorest 20%	Dummy variable equal to one if the individual has an income in the first income quintile, zero otherwise. Source: Global Findex Database.
Income-second 20%	Dummy variable equal to one if the individual has an income in the second income quintile, zero otherwise. Source: Global Findex Database.
Income-third 20%	Dummy variable equal to one if the individual has an income in the third income quintile, zero otherwise. Source: Global Findex Database.
Income-fourth 20%	Dummy variable equal to one if the individual has an income in the fourth income quintile, zero otherwise. Source: Global Findex Database.
Income-richest 20%	Dummy variable equal to one if the individual has an income in the fifth income quintile, zero otherwise. Source: Global Findex Database.
Log (GDP/capita)	Log of real Gross Domestic Product per capita. Source: World Development Indicators (World Bank).
Rule of law	Measures the perceptions of the extent to which people have confidence in and abide by the rules of society. Source: World Governance Indicators (WBG).
Log (population)	Log of the total population who are 15 years and above. Source: World Development Indicators (World Bank).
Catholic	Dummy variable equal to one if over 50% of inhabitants in a country are Catholics, zero otherwise. Source: The World Factbook (CIA).
Muslim	Dummy variable equal to one if over 50% of the inhabitants in a country are Muslim, zero otherwise. Source: The World Factbook (CIA).
Power distance	The extent to which individuals accept inequality. Source: Hofstede Insights.
Individualism	Measures the degree of interdependence a society maintains among its members. Source: Hofstede Insights.
Masculinity	Measures how distinct social gender roles are in a society. Source: Hofstede Insights.
Corruption	Corruption Perception Index. Source: Transparency International.
Trust	Index to measure trust. Source: La Porta et al. (1997).
Plough use	Measures the proportion of citizens with ancestors who traditionally used the plough in pre-industrial agriculture. Source: Alesina, Giuliano and Nunn (2013).

BOFIT Discussion Papers

A series devoted to academic studies by BOFIT economists and guest researchers. The focus is on works relevant for economic policy and economic developments in transition / emerging economies.

2019
- No 1 Çağatay Bircan and Orkun Saka: Lending cycles and real outcomes: Costs of political misalignment
- No 2 Lucy Chernykh, Denis Davydov and Jukka Sihvonen: Financial stability and public confidence in banks
- No 3 Yin-Wong Cheung and Shi He: Truths and myths about RMB misalignment: A meta-analysis
- No 4 Yuping Deng, Yanrui Wu, Helian Xu: Political connections and firm pollution behaviour: An empirical study
- No 5 Sophia Chen, Lev Ratnovski and Pi-Han Tsai: Credit and fiscal multipliers in China
- No 6 Alexander Kostrov and Mikhail Mamonov: The formation of hidden negative capital in banking: A product mismatch hypothesis
- No 7 Ning Cai, Jinlu Feng, Yong Liu, Hong Ru and Endong Yang: Government credit and trade war
- No 8 Michael Funke and Andrew Tsang: The direction and intensity of China's monetary policy conduct: A dynamic factor modelling approach
- No 9 Hamza Bennani: Does People's Bank of China communication matter? Evidence from stock market reaction
- No 10 Alexei Karas, William Pyle and Koen Schoors: Deposit insurance, market discipline and bank risk
- No 11 Gerard Roland and David Y. Yang: China's lost generation: Changes in beliefs and their intergenerational transmission
- No 12 Abel François, Sophie Panel and Laurent Weill: Are some dictators more attractive to foreign investors?
- No 13 Anna Pestova and Mikhail Mamonov: Should we care? The economic effects of financial sanctions on the Russian economy
- No 14 Haiyue Yu, Jin Cao and Shulong Kang: Fertility cost, intergenerational labor division, and female employment
- No 15 Max Breitenlechner and Riikka Nuutilainen: China's monetary policy and the loan market: How strong is the credit channel in China?
- No 16 Yiping Huang, Xiang Li and Chu Wang: What does peer-to-peer lending evidence say about the risk-taking channel of monetary policy?
- No 17 Heli Simola: Evaluating international impacts of China-specific shocks in an input-output framework
- No 18 Sris Chatterjee, Xian Gu, Iftekhar Hasan and Haitian Lu: Ownership structure and the cost of debt: Evidence from the Chinese corporate bond market
- No 19 Ke Song and Le Xia: Bilateral swap agreement and Renminbi settlement in cross-border trade
- No 20 Aaron Mehrotra, Richhild Moessner and Chang Shu: Interest rate spillovers from the United States: expectations, term premia and macro-financial vulnerabilities
- No 21 Zuzana Fungáčová, Eeva Kerola and Laurent Weill: Does experience of banking crises affect trust in banks?
- No 22 Mustafa Caglayan, Oleksandr Talavera and Wei Zhang: Herding behaviour in P2P lending markets
- No 23 Michael Funke, Xiang Li and Andrew Tsang: Monetary policy shocks and peer-to-peer lending in China
- No 24 Gayane Barseghyan: Sanctions and counter-sanctions: What did they do?
- No 25 Karlo Kauko: Benford's law and Chinese banks' non-performing loans

2020
- No 1 Chang Ma, John Rogers and Sili Zhou: The effect of the China connect
- No 2 Karlo Kauko: The vanishing interest income of Chinese banks
- No 3 Mariya Hake and Philipp Poyntner: Keeping up with the Novaks? Income distribution as a determinant of household debt in CESEE
- No 4 Risto Herrala and Fabrice Orlandi: Win-Win? Assessing the global impact of the Chinese economy
- No 5 Weijia Li, Gérard Roland and Yang Xie: Erosion of state power, corruption control, and political stability
- No 6 Ryan Banerjee, Boris Hofmann and Aaron Mehrotra: Corporate investment and the exchange rate: The financial channel
- No 7 Amanda Gregg and Steven Nafziger: Financing nascent industry: Leverage, politics, and performance in Imperial Russia
- No 8 Zuzana Fungáčová, Koen Schoors, Laura Solanko and Laurent Weill: Political cycles and bank lending in Russia
- No 9 Francis Osei-Tutu and Laurent Weill: Sex, language, and financial inclusion

BOFIT Discussion Papers
http://www.bofit.fi/en • email: bofit@bof.fi
ISSN 1456-4564 (print) // ISSN 1456-5889 (online)