

## CSAE Working Paper WPS/2016-12

# Gender-Oriented Languages and Female Labour Force Participation: Evidence from Sub-Saharan Africa †

Michelle Rao \*

2nd June, 2016 ‡

**Abstract:** Using the data from the Demographic and Health Surveys, this paper estimates the relationship between language and labour force participation of women in Sub-Saharan Africa. The results suggest that women who speak languages with stronger distinctions between masculine and feminine are less likely to participate in the labour force. This relationship holds both across and within countries, even after controlling for individual characteristics, religion and proxies for gender social norms related to ones ethnicity, such as historical use of the plough. The results suggest that language has a direct effect on preferences regarding labour market decisions, above and beyond gender norms arising from ethnicity and religion. These findings contribute to the growing literature on the relationship between socio-psychological factors and gender differences in economic outcomes.

**JEL Classification:** D03, J16, N37, 055

**Keywords:** Language; Identity; Culture; Gender social norms; Female labour force participation ; Sapir-Whorf hypothesis

---

†I thank Donna Harris, James Fenske, Douglas Gollin, Stefano Caria and Lutz Marten for their comments and suggestions.

\*University of Oxford: michelle.rao@gmail.com

‡A version of this paper was submitted as an extended essay for the MSc in Economics for Development at Oxford

# Gender-Oriented Languages and Female Labour Force Participation: Evidence from Sub-Saharan Africa

## 1. Introduction

Gender differences in labour market outcomes remain a significant hindrance to economic development (Duflo 2011). These differences persist over long periods of time, even despite improvements in economic growth and educational attainment (Blau and Kahn 2006). In explaining this trend, recent studies focus on the role of gender social norms and cultural beliefs that dictate behaviour for men and women in the labour force (Fortin 2005; Fernandez 2007).

This paper investigates the role of gender social norms on labour force participation through a consideration of language. In particular, languages are different in the degree to which they distinguish between biological sex in their noun classes. Whereas languages such as French have two noun classes, masculine and feminine, other languages, such as those from the Niger-Congo family, are known for complex noun classification systems with 10 or more noun classes (Dixon 1986).

I test the hypothesis that women speaking more gender-oriented languages are less likely to participate in the labour force. As argued by economists such as North (1991), language facilitates the transmission of culture across generations. The way a language marks gender “may influence the salience of gender in speakers’ mind (cognition) and/or may reflect the salience of gender distinctions in the culture,” (Boroditsky et al. 2013 p.65). Understanding the extent to which language influences thought, a term linguistics refer to as linguistic relativity, is therefore potentially important for our understanding of gender differences in labour market outcomes.

Using a range of linguistics sources, I follow Gay et al. (2013) in constructing a Number of Genders Intensity Index (NGII henceforth) equal to 1 if the language only has two noun classes, and 0 otherwise. Cross-linguistic studies on African languages show that languages with only two noun classes tend to be sex-based, whereas languages with a larger number of noun classes tend to be unrelated to sex distinctions (Di Garbo 2010). Though there are exceptions to this rule, NGII acts as a proxy indicator for the intensity by which the language marks gender.

In the first half of the paper, I run two sets of regressions of an indicator of female employment on NGII using individual level data on women in Sub-Saharan Africa from the Demographic and Health Surveys. The covariates for the first set of regressions are a rich set of controls for individual characteristics and country characteristics related to geography, history and contemporary economic environment. To ensure my results are not driven by unobserved effects across groups, I run a set of within-group regressions with fixed effect dummies for country, birth cohort, marital status and number of children. Finally, as I cannot fully rule out omitted variables bias, I follow Altonji, Elder and Taber (2000) and Oster (2013)

in quantifying the potential bias of the estimated treatment effect. My results consistently show that women speaking languages with only two noun classes are approximately 5 percentage points less likely to participate in the labour force. The magnitude of this relationship remains consistent and statistically significant, and is unaffected by changes in functional form. Furthermore, it does not seem to be driven by selection on unobservable characteristics.

In the second half of the paper, I investigate the potential mechanisms through which language may be affecting female labour supply. Firstly, I investigate whether the results are driven by gender social norms associated with those who speak gender oriented languages, using proxy controls such as religion and ancestral characteristics which have been shown to have effects on contemporary cultural beliefs (Alesina et al. 2011; Fenske, 2015; Nunn and Wantchekon, 2009). Secondly, I investigate the relationship between NGII and a woman’s number of children, educational attainment, and domestic violence, as these may be correlated with the net material benefits of working. Results from these sets of regressions provide suggestive evidence that the effect of language is not driven by differential material benefits to work, or from gender social norms arising from religion or from ancestral beliefs.

Broadly speaking, my results complement the existing empirical evidence for the importance of gender identity in labour market outcomes.<sup>1</sup> My paper can be seen as an indirect test of Akerlof and Kranton’s model of identity (2000, 2002, 2005). Through considering the role of language, I provide evidence in support of the importance of cognitive frameworks and psychological payoffs in explaining women’s roles in the labour market.

My findings also complement recent studies on language and economic outcomes. Chen (2013) finds that the way languages grammatically mark time is predictive of future-oriented behaviour. Most closely related to my study is work by Gay et al. (2013) who find that women speaking languages with stronger gender distinctions are less likely to participate in the labour market and political activities. I distinguish my paper from these works in two ways.

Firstly, both Chen (2013) and Gay et Al. (2013) rely heavily on the World Atlas of Linguistic Structures (WALS), a database covering 192 structural properties of world languages by a team of linguists. While this dataset covers a wide range of languages in countries worldwide, it lacks data for many African Languages. The results of these papers therefore have limited value for understanding the relationship of interest for countries in Sub-Saharan Africa. I address this by using a wider range of language resources. In total, I add an additional 47 African languages to my analysis that are not considered in Gay et al. (2003).

Secondly, estimation in Gay et al. (2013) relies primarily on cross-country variation. By integrating the Demographic and Health Surveys, I have a larger dataset with more detailed data on individual characteristics. This allows me to better account for unobservable characteristics across groups, through identifica-

<sup>1</sup>See Bertrand (2011) for a full review of the literature.

tion using variation for women within the same country, the same birth cohort, and with similar family structures.

Through focusing only on countries within Sub-Saharan Africa, my paper also provides an interesting complement to Gay et al.'s global study. Sub-Saharan Africa is one of the most linguistically diverse regions in the world, hosting more than 30% of the world's languages (Paul et al. 2009). Despite this diversity, many governments in African countries enforce the use of just one African or colonial language for medium of instruction and work. For this reason, linguistic issues in Sub-Saharan Africa are vastly different from those in the Western World. My focus on this region therefore provides a unique opportunity to understand the extent to which gender roles imposed by language survive interactions in long standing multi-lingual environments. My results suggest that the effect of language on labour market outcomes for women exists in this region, but has a noticeably smaller effect than that found in other countries.

## 2. Noun Classification Systems and Gender

Languages differ in the degree to which they make distinctions between sexes in their grammar. One common marker of biological sex is found in noun classification systems, "an obligatory grammatical system, where each noun chooses from a small number of possibilities" (Dixon 1986). The existence of a classification system is indicated by the presence of prefixes or obligatory articles that change according to a noun class.

Niger-Congo languages, one of the largest family of languages spoken in Sub-Saharan Africa, tend to have more than ten noun classes related to animate and inanimate objects and do not have indicators for biological sex. Swahili is one of these languages, and has a total of 18 noun classes. Two of these noun classes, Class 1 and Class 4,<sup>2</sup> are used for nouns referring to people and names of trees, respectively. As seen in the example below, the existence of a noun class system is evidenced by a corresponding change in the prefixes of the noun and verb.

- |    |                                    |                                   |
|----|------------------------------------|-----------------------------------|
| a) | <u>M</u> wanafunzi <u>a</u> nasoma | <u>M</u> ndimu <u>u</u> lianguka  |
|    | Student reads. CLASS 1. Singular   | Lime tree fell. CLASS 4. Singular |
|    | "The student is reading"           | "The lime tree fell"              |

**(Adapted from Ojiambo 2014)**

By contrast, the noun classification system in Somali, an Afroasiatic language spoken in Somalia, Ethiopia, Djibouti and Kenya, has only two noun classes – masculine and feminine. These are also evidenced outside of the noun itself, but by a corresponding change in the suffixes of the verb and noun.

<sup>2</sup>These are otherwise known as the "M-wa" and "M-mi" classes

- |    |   |                                      |
|----|---|--------------------------------------|
| b) | baabùur- <u>kii</u> waa <u>uu</u> yimid | náag- <u>tii</u> waa <u>ay</u> timid |
|    | Truck come. MASC. Singular              | Woman come. FEM. Singular            |
|    | “The truck came”                        | “The woman came”                     |

(Saeed 1999, p.56)

Following Gay et Al.’s methodology (2013), I therefore construct a Number of Genders Intensity Index (NGII) equal to 1 if the language only has two noun classes, and zero otherwise. As evidenced by Di Garbo’s (2013) cross-linguistic study, noun classification systems can be separated into two categories: languages with only two noun classes which tend to be sex based and languages with a large number of noun classes that do not draw distinctions between the sexes.<sup>3</sup> Therefore, though imperfect, NGII acts a proxy for the intensity by which the language draws distinctions between biological sexes.

### 3. Hypothesis and Mechanisms

The main hypothesis of the paper is that women speaking more gender-oriented languages are less likely to participate in the labour force. This hypothesis stems from two potential effects of language on saliency and gender social norms. In this section, I illustrate these effects through building a simple model of binary work choice.

Critically, I follow Akerlof and Kranton (2000, 2002, 2005) in assuming that a woman experiences cognitive dissonance if their actions deviate from those prescribed by societal gender norms. A woman’s utility therefore comprises two parts: a material payoff, dependent on consumption and leisure, and a psychological payoff, that applies to women who work. I assume for simplicity that the gender prescribed norm for all women is to not work. Therefore, a psychological disutility is felt by woman  $i$  only if they choose to work.

Letting the utility function be additively separable in psychological and material payoffs, a woman’s utility given their work status is:

$$(1) \quad U_i^W = f(c_i^W, l_i^W) - W \cdot g(I^*, \rho_i)$$

where  $W$  is a binary indicator for work status. The material payoff from working is captured by  $f(c_i^W, l_i^W)$ , a function of consumption  $c_i^W$ , and leisure  $l_i^W$ . The magnitude of the psychological cost of working is represented by  $g(I^*, \rho_i)$ , where  $g(I^*, \rho_i) \geq 0$ .

$g(I^*, \rho_i)$  varies by individual depending on two factors. It firstly depends on  $I^*$ , the magnitude of the common disutility from working that is shared within a cultural group (whether this is by country, ethnic group, or otherwise). As

<sup>3</sup>There are also instances where languages have a mix of sex-based and non sex-based gender assignment systems. In his study of African language typology, however, Heine (1982) argues that these are very rare.

$I^*$  is not perfectly observable, the individual specific psychological payoff further depends on  $\rho_i$ , a term that reflects the degree to which the woman correctly values  $I^*$ . This can, for example, be interpreted as the saliency of gender social norms. I assume the following:

$$(2) \quad g(I^*, \rho_i) = \rho_i I^*, \quad \rho_i \in [0, 1] \forall i, \quad I^* \geq 0$$

Therefore, a woman will choose to work if:

$$(3) \quad f(c_i^{W=1}, l_i^{W=1}) - f(c_i^{W=0}, l_i^{W=0}) \geq \rho_i I^*$$

That is, if the net material benefit of working is greater than the psychological cost to working.

Assuming characteristics of one's language do not differentially affect the net material benefit of working, the results from this model give two main mechanisms through which speaking a language with stronger sex-distinctions may reduce labour force participation for women.

**Mechanism 1 – Saliency:** Languages that make stronger distinctions between masculine and feminine increase the saliency,  $\rho_i$ , of gender social norms. As seen in section two, evidence for noun classes is found *outside* of the noun itself. If the language has a noun classification system which is sex based this increases the psychological disutility of working as individuals become more aware of  $I^*$ . That is, by marking gender in linguistics, “social gender is ‘called up’ [...] mak[ing] it difficult for a speaker to ignore gender” (Ecker and McConnell-Ginet, 2013, p.86). Therefore women speaking NGII = 1 languages are less likely to work.

**Mechanism 2 – Gender Social Norms:** Women who speak languages with strong sex-distinctions tend to be from societies with stronger associated gender social norms, and hence a higher  $I^*$ . As the assignment of nouns into noun classes is often rooted in semantics (Aikhenvald 2003), the classification of nouns into sex-based categories could reflect the gender norms inherent to ancestors of those speaking the language. The effect of language on current economic outcomes would therefore be evidence of inter-generational transmission of gender norms that are deeply rooted in societies that speak sex-oriented languages.

## 4. Data

### 4.1. Demographic and Health Surveys

I use the data on women living in Sub-Saharan Africa from the Demographic and Health Surveys (DHS) Program, a set of surveys of nationally representative women over the age of 15. Use of the DHS gives me an advantage over previous studies of language and economic outcomes for the following reasons. Firstly, it provides me with comprehensive indicators for language, employment, and

familial characteristics at the individual level. Secondly, the datasets are very large, which allows me to use methods that are more statistically demanding on the data, such as dummies for country and birth cohorts, without sacrificing precision.

I use the most recent survey wave that has data on language spoken and data related to a women’s status and empowerment, including indicators on employment, education, and marriage. After merging the data, I have a total of 391,701 observations across 25 countries in Sub-Saharan Africa.<sup>4</sup> Summary statistics for main variables and individual characteristics of the subsample are found in **Table 1**.

Table 1—: Individual level characteristics

	(1) Observations	(2) Mean (Std. Error)
NGII	267,659	0.36 (0.48)
Employed	267,659	0.60 (0.49)
Service	267,659	0.046 (0.21)
Agriculture	267,659	0.25 (0.43)
Age	285,188	28.7 (9.74)
Years of Education	285,136	5.25 (4.89)
Children	284,463	2.86 (2.85)
Wealth Index	285,188	3.13 (1.43)
Marital Status	285,188	0.63 (0.48)
Urban Status	285,188	0.38 (0.49)
Christian	285,188	0.59 (0.49)
Muslim	285,188	0.27 (0.45)

### Main Variables:

**Primary language spoken.** There are three indicators related to language that appear in the DHS surveys: native language, language spoken at home, and language of the interview. When the first two indicators are available, I use them

<sup>4</sup>The survey rounds used include: BF6, BJ6, CD6, CG5, CI6, CM6, ET6, GH6, GN6, KE6, KM6, LS5, ML5, MW5, NG6, NI6, NM6, RW4, SL6, SN6, SZ5, TG6, UG6, ZM6, ZW6

as the main indicator for language. Otherwise, I use language of the interview as a proxy for primary language spoken. When it is reported an interpreter was needed to conduct the interview, I exclude the observation from my sample. This leaves a total of 73 languages.

**Women’s status and empowerment.** The DHS conducts a set of surveys with questions related to a woman’s status and empowerment. Based on these questions, I define a set of dummy variables of interest. “*Employed*”, is a dummy variable that indicates whether or not the individual considers herself as employed at the time the questionnaire was conducted. Based on the DHS grouping of occupations, I define indicators for the individual’s occupation, such as “*Service*” and “*Agriculture*”.

#### 4.2. Classification of Languages

To classify languages into their respective NGII categories, I use a range of data sources to find the number of noun classes of each language in my DHS dataset. The main data source used for this is the World Atlas of Linguistic Structures Online (WALS), a database covering 192 structural properties of world languages by a team of linguists. Within this database, I use the chapter on “Number of Genders” by Corbett (2013) who categorizes each language into their respective number of noun classes. Matching the WALS languages with the DHS gives the noun class systems for 14 languages in the DHS.

For the remaining 59 languages, I use a variety of sources to classify the languages. I use AboutWorldLanguages.com and Ethnologue, online resources that provide information on the history and grammar of world languages. These databases allow me to match an additional 15 languages on the DHS. The rest of the languages are classified manually using a range of African language grammar books (see appendix for citations). These methods allowed me to classify a total of 60 languages in my dataset, leaving a total of 267,659 observations. A complete list of the language name, language family, number of noun classes and information source can be found in the appendix.

### 5. Overview of Empirical Strategy

The main empirical model used in this paper is the Linear Probability Model (LPM), equivalent to ordinary least squares with a binary dependent variable. Whilst other functional forms, such as Logit and Probit, are often used for binary dependent variables, I use LPM due to ease of interpretation of marginal effects and flexibility of specification (Angrist and Pischke 2008). I show that my results are robust to different functional form and distributional assumptions in the robustness checks.

Using the LPM, the conditional probability of an individual being employed given their characteristics is assumed to be linear such that the expected value of

employment (conditional on NGII and covariates) is given by:

$$(4) \quad E(y_{ilc}|NGII_l, \mathbf{X}_{ilc}, \mathbf{Z}_{lc}; \alpha, \beta, \boldsymbol{\tau}', \boldsymbol{\gamma}') = \alpha + \beta NGII_l + \boldsymbol{\tau}'\mathbf{X}_{ilc} + \boldsymbol{\gamma}'\mathbf{Z}_{lc} + \epsilon_{ilc}$$

where  $y_{ilc}$  is a binary variable, indicating whether individual  $i$  with language  $l$  and living in country  $c$  is employed. Our main variable of interest is  $NGII_l$  an indicator for whether the language spoken by the individual has only two noun classes.  $\mathbf{X}_{ilc}$  is a  $m \times 1$  column vector of individual level characteristics, and  $\mathbf{Z}_{lc}$  is an  $n \times 1$  column vector of country specific controls.

The main econometric model can therefore be expressed as follows:

$$(5) \quad y_{ilc} = \alpha + \beta NGII_l + \boldsymbol{\tau}'\mathbf{X}_{ilc} + \boldsymbol{\gamma}'\mathbf{Z}_{lc} + \epsilon_{ilc}$$

The fitted value from the model gives the predicted probability of an individual being employed given her characteristics. Standard errors are clustered by language. This allows me to account for (1) the inefficiency that arises from heteroskedastic standard errors inherent in LPM models, and (2) any correlation in labour force participation outcomes within language groups. The latter is particularly important given our context: as NGII only varies at the level of the language, we need to account for intra-class correlation so as not to overestimate precision (Moulton 1986).

**Identification Problem:** Similar to standard OLS, the LPM produces consistent and unbiased estimates of the true relationship between NGII and employment only if the zero conditional mean assumption holds:

$$(6) \quad E(\epsilon_{ilc}|NGII_l; \mathbf{X}_{ilc}, \mathbf{Z}_{lc}) = 0$$

Common threats to this are reverse causality and omitted variables bias. In linguistics, the gender orientation of a language is considered to be one of the most mature and stable features of language, formed prior to long evolutionary chains (Dahl 2004; Wichmann and Holmon 2009). Therefore, I assume there is no reverse causality, as the gender orientation of a language is likely to be exogenous.

I am however, concerned with omitted variables bias, which arises from omitting covariates that may be correlated with NGII and with labour force participation. For example, an individual's adoption of language may have colonial origins. This is particularly relevant for African countries, with official languages and language of educational instruction often being the former colonial language. At the same time, colonial origins are shown to have long run effects on economic growth and development, whether this is through their effect on institutions or socioeconomic structures (Engerman and Sokoloff 1997, 2002; La Porta et al. 1998; Acemoglu et al. 2001, 2002). Omission of colonial origins in the main specification would therefore be likely to lead to an inconsistent estimate of the true relationship between NGII and employment. Whilst I am able to control for some of these

omitted variables others are unobserved and cannot be appropriately accounted for.

The remainder of this paper seeks to minimize this bias in order to get closer to the direct causal relationship between gender-oriented languages and labour force participation. In section 6.1, I use a linear probability model firstly with controls, and secondly with fixed-effect dummies, to account for selection on country or individual-level characteristics that may be driving the results. As I cannot rule out omitted variables bias, I then quantify the potential bias driven by unobservable heterogeneity in section 6.2. In section 7, I control for contemporary religious beliefs and ancestral characteristics. These variables act as proxies for gender beliefs and values, allowing me to account for cultural beliefs that may be associated with those who speak NGII languages. This enables me to get closer to distinguishing between the mechanisms through which language may be affecting work choice for women.

## 6. Main Results

### 6.1. Baseline Results

#### **Individual and Country level controls:**

My first set of regressions, as specified in equation 5 includes a rich set of individual level and country specific controls,  $X_{ilc}$  and  $Z_{lc}$ , that may be correlated with both the error term and NGII.

**History Controls:** As previously explained, colonial origins are likely to be correlated with both employment and adoption of language. Common colonial languages such as French and Portuguese only have two noun classes. If having a French or Portuguese colonizer also led to stronger institutions and higher employment, then omission of these indicators would lead to an overestimate of the true effect of NGII. Toward this end, I include dummy variables for colonial origin in my main specification. I also control for ethnic fractionalization and historical country-level slave exports, which Alesina and La Ferrara (2005) and Nunn and Wantchekon (2009), respectively show to be negatively correlated with economic development today. These controls allow me to additionally account for adoption of languages that may be associated with historical interactions between different cultures.

**Geography controls:** I follow Nunn (2008) in controlling for absolute latitude, the logarithm of coastline divided by the land area, and weather indicators, including average minimum monthly rainfall, maximum humidity, and temperature. These allow me to account for the effect of geography on long run economic development that has been found in the literature (Bloom et al. 1998; Sachs et al, 2001)– whether this is through prevalence of infectious diseases, agricultural productivity or other channels.

**Contemporary controls:** I follow La Porta et al. (2008) and Chen (2013) in their cross-country regression methodology in controlling for log GDP per capita,

growth of GDP per capita and the unemployment rate. These serve as indicators for the economic environment of the country, which in turn affects ease of job search. I also control for trade openness, which Özler and Nilüfer (1995) have shown to be correlated with female labour force participation.

**Individual controls:** At the individual level, I follow Alesina et al. (2013) in controlling for educational attainment, marital status, urban status, age and age squared. Controlling for age squared allows me to account for potential age effects that would otherwise be omitted in the specification. I further control for wealth status<sup>5</sup> and number of children at home as these may affect the net material benefit to working.

Estimates of equation 5 are reported in columns 1-4 of Table 2. Regressions (1) and (2) use only exogenous controls for identification. Regression (1) reports the relationship between NGII and employment controlling for age and age squared. This coefficient estimate summarizes the relationship between NGII and participation in the labour force: on average, women speaking languages with only two noun classes are 4.63 percentage points less likely to work, though this is not significant at the 10% level. Regression (2) adds in controls for history and geography to ensure we are not attributing the relationship between NGII and language to cross-country differences in fixed characteristics pertaining to an individual's country of residence. After accounting for fixed country level characteristics, the predicted relationship between NGII and employment increases in magnitude, and is significant at 5%.

As this result may be confounded by market conditions specific to the country or to the individual, regression (3) adds the full set of contemporary controls related to a country's economic conditions. This allows me to partial out any correlation between NGII and employment that may be driven by labour market conditions affecting the availability and ease of securing employment. Addition of contemporary country level controls increases the predicted negative effect of NGII to 5.72 percentage points.

Regression (4) adds in the entire set of individual level controls that may affect the net material benefit of working: these include variables such as marriage, number of children, urban status and wealth. This gives us the main result of the paper: a woman speaking a language with only two noun classes is 5.15 percentage points likely to participate in the labour force.

Note here that the full set of individual characteristics and contemporary controls are potentially outcomes of gender-oriented languages. Therefore, they could be seen as "bad controls" which, although controlling for omitted variables bias, would also introduce its own selection bias (Angrist 2008). As the coefficients do

<sup>5</sup>Wealth Status is a DHS indicator for standard of living, measured by ownership and access to assets and basic facilities. The index is on a scale of 1-5, in ascending order of wealth. My specification uses dummy variables for each category, excluding the first.

Table 2—: Main Results

	(1) Employed	(2) Employed	(3) Employed	(4) Employed	(5) Employed	(6) Employed	(7) Employed
NGII	-0.0463 (0.0288)	-0.0567** (0.0220)	-0.0572** (0.0218)	-0.0515** (0.0208)	-0.0585** (0.0285)	-0.0582** (0.0284)	-0.0580** (0.0284)
Individual Controls	Age and Age Squared	Age and Age Squared	Age and Age Squared	YES	YES	YES	YES
History Controls	NO	YES	YES	YES	NO	NO	NO
Geography Controls	NO	YES	YES	YES	NO	NO	NO
Contemporary Controls	NO	NO	YES	YES	NO	NO	NO
Country FE	NO	NO	NO	NO	YES	YES	YES
Birth Cohort FE	NO	NO	NO	NO	NO	YES	YES
Marital Status x Children FE	NO	NO	NO	NO	NO	NO	YES
Observations	267,659	267,659	267,659	266,885	266,885	266,885	267,590
R-squared	0.112	0.156	0.157	0.160	0.173	0.174	0.177

*Note:* Standard errors clustered at level of the language are reported in parentheses. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%. Description of the set of controls are as follows: Individual controls are age, age squared, educational attainment, number of children and marital, urban and wealth status; History controls include ethnic fractionalization (from Alesina and La Ferrara 2005), historical country-level slave exports (from Nunn and Wantchekon 2009) and dummy variables for colonial origin (including indicators for Britain, France, Portugal, Belgium, Spain, UN and Italy); Geography controls are from Nunn (2008), including absolute latitude, logarithm of coastline divided by land area, average minimum monthly rainfall, maximum of monthly afternoon humidity and minimum of average monthly temperature; Contemporary controls include log of GDP per capita, growth of GDP per capita, the unemployment rate and trade openness.

not change largely by adding these controls (comparison of regressions 2-4), we can be relatively confident that this form of selection is not driving our results.

### Within-Group Regressions:

Despite the rich set of controls at the country and individual level, the results from regressions (1)-(4) could still be driven by unobserved differences across countries or in norms and values across generations and family structures. Toward this end, I conduct a series of regressions with individual level controls and fixed effects, at the level of the country and birth cohort. I then follow Chen (2013) in including fixed effects for marital status and number of children. This allows me to estimate the effect of NGII solely based on variation within groups. I therefore rule out any identification that may be driven by unobservable characteristics across countries, generations and family structures.

I estimate the following:

$$(7) \quad y_{ilc} = \alpha + \beta NGII_l + \tau' \mathbf{X}_{ilc} + \sum_{c=1}^C \delta_c \text{country}_c + \sum_{i=1}^I \phi_i \text{birthcohort}_i + \sum_{i=1}^I \pi_i \text{maritalstatus}_i \cdot \text{children}_i + \epsilon_{ilc}$$

where  $\mathbf{X}_{ilc}$  again includes the set of individual specific characteristics. Unless otherwise stated for reasons of multicollinearity, the set of controls included are the same set of controls as those previously defined. Here,  $\text{country}_c$  is a dummy variable set to 1 if the individual lives in country  $c$ .  $\text{birthcohort}_i$  are indicators for the individual's year of birth, defined in five year intervals.  $\text{maritalstatus}_i$  are a set of indicators for the individual's marital group, as defined by the DHS.<sup>6</sup> Lastly,  $\text{children}_i$  indicates an individual's number of children, in groups of five. I refer to the interacted fixed effects of marital status and number of children as "family structure" fixed effects.

Two potential problems may occur when using fixed effects of this kind. Firstly, whilst this methodology allows me to control for a richer set of omitted variables, it is also more demanding on the data. In the most demanding specification with country, birth cohort and family structure fixed effects, identification relies heavily on variation within these subgroups rather than large variations across groups. This means that we may lose power and precision. Secondly, if there is classical measurement error in our data, fixed effects can worsen attenuation bias as the signal to noise ratio is likely to be lower for within-group variation.

Results for the within-group regressions are reported in columns 5-7. Regres-

<sup>6</sup>There are 5 marital groups reported in the DHS survey: never in union; married; living with partner; widowed; divorced; and no longer together/separated

sion (5) controls only for country fixed effects, using only within country variation. Regression (6) identifies the relationship between NGII and labour force participation by comparing only individuals within the same country and born in the same five-year span. This estimate allows the conditional mean of the distribution to change across countries and across birth cohorts.

Finally, in regression (7), I add fixed effects for marital status and number of children, interacted. These allow me to account for a woman's family structure, which may be affecting the material benefits of working. Again, the estimates for the relationship between NGII and labour force participation do not change largely in magnitude and are still significant at the 5% level. Note here that similar to the full set of controls for individual characteristics, these family structure fixed effects could be interpreted as outcomes of language. The estimated relationship here shows that, even going above and beyond the potential effect of language on outcomes related to marriage and children, the relationship between NGII and employment still holds.

Importantly, estimation in regressions 5-7 allows me to control for any unobservable characteristics across birth cohorts, countries or family structures that may have been driving the results in regressions 1-4. Despite the standard errors being slightly larger, the results are still significant at the 5% level. The increase in magnitude of the coefficients suggests measurement error is not biasing the estimated relationship.

Depending on the specification, the size of the coefficient ranges from  $-0.0515$  to  $-0.0580$ . At the lower bound, my results imply that changing from a NGII = 0 to a NGII = 1 language is associated with a fall in employment that is equal to 8.27% of the average and 10.5% of one standard deviation of the employment rate. Therefore, though not large, when put in context the estimated effect has a meaningful effect on employment.

## 6.2. Selection on Unobservables

There is still the possibility that the estimated effect is driven by unobserved omitted variables, such as institutional quality, that are not fully accounted for by the set of controls. Toward this end, I use tests proposed by Altonji, Elder and Taber (2000) and Oster (2013) to assess the severity of the potential bias due to unobservable heterogeneity.

The basis of this test is the assumption of proportional selection which, in the context of this paper, can be defined as follows:

$$\delta \frac{Cov(Observed, NGII)}{Var(Observed)} = \frac{Cov(Unobserved, NGII)}{Var(Unobserved)}$$

where  $\delta$  is the coefficient of proportionality. That is, the relationship between the unobserved variables and NGII (the treatment) is proportionate to the relationship between the observed controls and NGII. Given a known  $\delta$ , Oster (2013) then shows that we are able to account for the selection due to unobservables, and

therefore derive a consistent and unbiased estimator for the treatment effect. Note however, that this method is unable to account for “unobserved controls which *do not* share covariance properties with the observed controls” (Oster 2013, p.9).

In practice, we cannot calculate  $\delta$ . However, we can estimate the unbiased treatment effect given different levels of  $\delta$ . I follow Altonji, Elder and Taber (2005) and Oster (2013) in assuming  $\delta \in [0, 1]$ : as the controls are chosen precisely to partial out correlation driven by unobservables, Altonji et al. (2005) argue an upper bound of 1 is suitable for applied analysis. Using this assumption, I calculate the estimated unbiased treatment effect under the assumption that  $\delta = 0$  and  $\delta = 1$ , where  $\delta = 0$  is the most optimistic assumption which assumes there is no selection on unobservables (and hence no omitted variables bias), and  $\delta = 1$  provides the most conservative estimator, which assumes

$$\frac{Cov(Observed, NGII)}{Var(Observed)} = \frac{Cov(Unobserved, NGII)}{Var(Unobserved)}$$

Table 3—: Coefficient Bounds

	(1) Employed	(2) Employed	(3) Employed
NGII	-0.0515** (0.0208)	-0.0585** (0.0285)	-0.0580** (0.0284)
Coefficient Bounds	(-0.515, -0.0128)	(-0.585, -0.0570)	(-0.580, -0.0541)
Controls	Complete Set of Controls	Country Fixed Effects & Individual Controls	Country, Birth Cohort and Family Structure FE, Age and Age Squared
Observations	266,885	266,885	267,590
R-squared	0.160	0.173	0.177

*Note:* Standard errors clustered at level of the language are reported in parentheses. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%. Coefficient bounds are calculated using the PSACALC command generated by Oster (2013). “Complete set of controls” refers to the full set of individual, history, geography and contemporary controls described in [Table 2](#). “Individual Controls” are also listed in [Table 2](#).

As seen by the coefficient bounds in [Table 3](#) my results do not seem to be driven by selection on unobservables that are systematically correlated with the observable characteristics. Bias due to selection on unobservables is particularly high with the full set of individual level and country level controls: here, the upper bound on the estimated effect is reduced to -0.0128 though it is not elimi-

nated. By contrast, the coefficient bounds on regressions (2) and (3) suggest that even under the most conservative assumption of  $\delta = 1$ , the estimated unbiased treatment effect remains statistically significant and similar in magnitude. This suggests that, although the size of the estimated treatment effect is similar between specifications (1), (2) and (3), regressions with fixed effects (as in (2) and (3)) are able to better account for omitted variables.

## 7. Investigating the Mechanisms

The results thus far do not allow me to distinguish between the potential mechanisms through which gender-oriented language may be affecting work choice for women. In this section, I address this by directly controlling for contemporary and ancestral characteristics that are shown to be associated with gender norms and beliefs in the literature.

### 7.1. Religion

In **Table 4** I run a regression of employment on NGII with additional controls for religion. I do this by using indicators for an individual being Muslim and for being Christian, the two most prominent religions in my sample of countries.

Regressions (1) and (2) control for *Christian* and *Muslim* separately. The coefficient estimates show that religion has a large and statistically significant effect on female labour supply. Regression (2), for example, shows that women who are Muslim are 6.03 percentage points less likely to work. This effect is larger than the estimated impact of NGII, but notably, the effect of language is not eliminated. After including both indicators in regression (3), the effect of NGII reduces to -0.0482. The coefficient is only significant at 10%, though this does not seem to be driven by an increase in noise.

Rather, the reduction in magnitude and significance is reflective of two factors. First, consistent with the literature (Lehrer 1995, 2004; Campante and Drott 2013), gender beliefs arising from the Islam religion have a negative effect on female labour supply. Second, being Muslim is likely to be positively correlated with speaking a language with only two noun classes. For example, Arabic is a Nilo-Saharan language with only two noun classes, and is most commonly spoken by Muslims worldwide. As *Muslim* is negatively correlated with NGII and *Employed*, its addition into the regression reduces the magnitude of the estimated coefficient.

Despite this, the coefficient on NGII is still significant at 10%. Though the effect of NGII on labour force participation is still smaller than that of *Muslim*, the role of NGII does not go away. Comparing this to the same regression without controls for religion in **Table 2** (coefficient: -0.0585; standard error: 0.0285), the results suggest that just under 20 percent of the relationship between NGII and employment can be explained by beliefs arising from religious beliefs. This

demonstrates that the effect of language on labour supply is not driven wholly by differential gender beliefs arising from religion.

Table 4—: Controlling for Religion

	(1) Employed	(2) Employed	(3) Employed
NGII	-0.0517* (0.0277)	-0.0481* (0.0277)	-0.0482* (0.0278)
Christian	0.0475*** (0.0163)		0.00966 (0.0170)
Muslim		-0.0603*** (0.0222)	-0.0517* (0.0278)
Country Fixed Effects	YES	YES	YES
Individual Controls	YES	YES	YES
Observations	266,885	266,885	266,885
R-squared	0.174	0.174	0.174

*Note:* Standard errors clustered at level of the language are reported in parentheses. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%. The full set of individual controls are described under [Table 2](#).

## 7.2. Ancestral Characteristics

To further distinguish between the mechanisms, I follow the growing literature in economic history on culture and control for ancestral characteristics that are potentially related to views and beliefs on the role of women in society today. If the relationship between NGII and employment remains robust to these controls, the results would provide suggestive evidence of mechanism 1, saliency.

For this analysis, I merge DHS ethnicity to that from Murdock’s Ethnographic Atlas. Not all ethnicity names had a direct matching of ethnicity name with the Ethnographic Atlas. Some of women were merely not asked about their ethnicity. Of those that were asked, many responses – such as “African” or “Other” were uninformative; many ethnicities were also simply not included in the Ethnographic Atlas. This leaves me with 174,884 observations for the analysis. A regression of employment on NGII with individual level controls and country level dummies gives an estimated coefficient, -0.0583, that is significant at the 10%. This suggests that, albeit loss of precision potentially due to reduction in sample size, the use of the subset of women for the analysis in this section is relatively representative of the main DHS sample.

In [Table 5](#), I control for a set of ancestral characteristics that are shown to be

correlated with gender social norms in the literature. As these characteristics have been shown to affect other individual outcomes relating to factors such as marital status, education and number of children, I only include exogenous individual controls for this analysis – age and age squared. Column 1 of **Table 5** reports the baseline regression with controls for age, age squared, and country level fixed effects.

Table 5—: Controlling for Ancestral Characteristics

	(1) Employed	(2) Employed	(3) Employed	(4) Employed	(5) Employed	(6) Employed
NGII	-0.0668*	-0.0678*	-0.0639*	-0.0666*	-0.0664**	-0.0666**
	(0.0331)	(0.0353)	(0.0349)	(0.0333)	(0.0329)	(0.0324)
Plough		0.0229 (0.0371)				
Agricultural Suitability			-0.0222 (0.0333)			
Subsistence Fishing or Hunting				-0.0141 (0.0286)		
Subsistence Agriculture					0.00589 (0.0364)	
Bride Price						0.118*** (0.0253)
Individual Controls	Age and Age Squared	Age and Age Squared	Age and Age Squared	Age and Age Squared	Age and Age Squared	Age and Age Squared
Country Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	174,923	166,762	173,986	171,837	174,923	174,923
R-squared	0.169	0.166	0.168	0.168	0.169	0.170

*Note:* Standard errors clustered at level of the language are reported in parentheses. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%. Plough, Subsistence Fishing or Hunting, Subsistence Agriculture and Bride Price are variables generated from Murdock’s Ethnographic Atlas. Agricultural Suitability is a measure generated by Alesina et al. (2013).

I firstly control for historical use of the plough. Alesina et al. (2013) find that historical use of the plough is predictive of contemporary gender social norms and therefore, of female labour force participation today. They argue that this is because cultivation using the plow is more laborious and demanding for physical

strength. Therefore, “societies that traditionally practiced plough culture developed a specialization of production along gender lines” (Alesina et al. 2013, p1). This effect is transmitted across generations, and so women of ethnic groups from these societies are less likely to participate in political and economic life today.

Regressions (2) and (3) therefore control for historical use of the plow in two ways: firstly, by using a dummy indicator for whether plow cultivation was present in the ethnic group’s society; secondly, by using an indicator data generated by Alesina et al. (2013) on the historical suitability of land and soil for growing crops that require plough cultivation. In both specifications, the magnitudes of the estimates are not largely affected, though the significance of NGII reduces to 10%, likely due to the increase in noise. The results suggest that the effect of gender-oriented language on female labour force participation is independent of gender attitudes driven by historical differences in agriculture.

In regressions (4) and (5) I control for the ancestral mode of subsistence, which Michalopoulos and Putterman (2014) show to be correlated with contemporary education and wealth. Finally, in regression (6) I control an indicator of historical requirement of bride price that Alesina et al. (2015) show are associated with less contemporary violence towards women today. The coefficients remain stable in magnitude and are statistically significant at the 5% or 10% level.

The results suggest that the ancestral characteristics are not part of the causal channel through which NGII affects female labour force participation: therefore, the effect of NGII on labour choice is likely to be independent of gender norms arising from these ancestral characteristics.

## 8. Gender Segregation in the Labour Market

Having established a relationship between NGII and labour force participation for women, I now turn to a woman’s *choice* of occupation to understand the degree to which NGII can explain differential labour market outcomes for women. I restrict my sample to only women who are employed and define two variables – service and agriculture – equal to 1 when the individual reports working in the service and agricultural sector, respectively.

The results suggest that there is some degree of horizontal gender segregation in employment that can be explained by NGII: women who are working and speak gender-oriented languages are 9.96 percentage points less likely to work in agriculture. As the results could potentially be explained by the suitability of the land to agriculture, I include a measure of agricultural suitability at the ethnicity level constructed by Alesina et al. (2013) in regressions (3) and (4). The estimated effect for *agriculture* is very similar, and remains statistically significant at 5%. No similar effect is found for services. This is suggestive of the effect of language on work-choice preferences.

Table 6—: Gender Segregation in Employment: Agriculture and Services

	(1) Agriculture	(2) Service	(3) Agriculture	(4) Service
NGII	-0.0996** (0.0464)	0.0000783 (0.0220)	-0.0941** (0.0422)	0.00225 (0.0207)
Agricultural Suitability			-0.0410 (0.0391)	-0.0161 (0.0285)
Individual Controls	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Observations	174,884	174,884	174,884	174,884
R-squared	0.250	0.133	0.251	0.133

*Note:* Standard errors clustered at level of the language are reported in parentheses. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%. The full set of individual controls are described under [Table 2](#).

## 9. Other Dependent Variables

So far, I have assumed that gender-oriented languages are affecting labour force participation through their effect on psychological payoffs. However, as illustrated in equation 3, one potential interpretation of the results is that the net material benefit of working is different for women who speak more gender-oriented languages, versus those who do not. To investigate whether this is the case, I now consider whether the effect of NGII extends to other household outcomes which may differentially affect the material opportunity cost of working for NGII speakers.

I consider three dependent variables. First, I consider the woman's number of children as this may be associated with higher costs to working, for example due to higher childcare costs. Secondly, I consider years of education as this is likely to be associated with higher returns to working. Lastly, I investigate the relationship between NGII and *violence*, where the latter is an indicator for whether the woman reported 'yes' to experiencing sexual or any form of physical violence from her spouse. According to the male backlash theory (Anderson 1997; Gartner and Macmillan 1999), men may resort to domestic violence as a means of asserting power if they feel threatened by their wives working. Therefore, domestic violence acts as a proxy for constraints to working imposed by husbands, which may be reducing female labour supply.

The effect of gender-oriented languages on these dependent variables is reported in [Table 7](#). As seen from regressions (1) to (3) the effect of NGII does not seem to extend to these dependent variables. This suggests that the effect of NGII on employment is not through differential material benefits of working.

Table 7—: Other Dependent Variables

	(1) Number of Children	(2) Years of education	(3) Violence ever
NGII	-0.2.07 (0.246)	1.279 (1.159)	-0.435 (0.474)
Individual Controls Country FE	Age and Religion YES	Age and Religion YES	Age and Religion YES
Observations	207,663	208.213	158,102

*Note:* Standard errors clustered at level of the language are reported in parentheses. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%. I control for religion by including indicators for the individual being Muslim and for being Christian; Age is controlled for by using age and age squared

## 10. Robustness Checks

### 10.1. Functional Form and Standard Errors

#### Functional Form

I report my main results using the logit specification. In addition to showing that my results are robust to different functional forms, this allows me to compare the estimated coefficient in this paper to that from Gay et al. (2013).

Using the logit model, the conditional probability of working given individual and country level characteristics is given by:

$$(8) \quad Pr(y_{ilc} = 1 | NGII_l, \mathbf{X}_{ilc}, \mathbf{Z}_{lc}) = \Lambda(\beta_0 + \beta_1 NGII_l + \boldsymbol{\tau} \mathbf{X}_{ilc} + \boldsymbol{\tau} \mathbf{Z}_{lc})$$

where the error term is assumed to be i.i.d distributed by the logistic CDF.

Regression (1) in **Table 8** reports the estimated marginal effects using the logit specification with the full set of country level and individual level controls. Despite changes in the functional form, the estimated coefficient remains statistically significant at 5%. Using a t-test to compare our estimated odds ratio of 0.771 with Gay et al.'s (2013) finding of 0.542 (standard error of 0.114), we can reject the null hypothesis that the two estimated effects are equal. Though the dataset used between the two studies are different, this provides suggestive evidence that the effect of gender marking is not as large in my subset of countries.

Table 8—: Robustness Checks

	(1) Logistic (Marginal Effect)	(2) Logistic (Odds Ratio)	(3) LPM	(4) Logistic (Marginal Effect)
	Employed	Employed	Employed	Employed
NGII	-0.0619** (0.103)	0.771** (0.0797)	-0.0515** (0.0213)	-0.0619** (0.0248)
Clustering	Language	Language	Language & Country	Language & Country
Individual Controls	YES	YES	YES	YES
History Controls	YES	YES	YES	YES
Geography Controls	YES	YES	YES	YES
Contemporary Controls	YES	YES	YES	YES
Observations	266,885	266,885	266,885	266,885

*Note:* Standard errors clustered at level specified. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%. Column headers indicate the functional form used for the specification. A full description of the full set of individual, history, geography and contemporary controls can be found under **Table 2**.

### Two-way clustering

In my main specifications, I account for potential within language correlation by clustering standard errors at the level of the language. However the effect of gender-oriented language on labour force participation could also be correlated within each country, particularly among those that speak the same language in the same country. I account for this possibility using two-way clustering in both the linear probability model and the logit specifications in columns (3) and (4). Even after controlling for both within country and language correlation, the relationship remains significant at 5%.

#### 10.2. Variation in NGII by country

In this section, I run my main specification for subsets of my DHS sample with different levels of NGII variation. In **Figure 1**, I plot the relationship between share of women speaking gender-oriented languages and share of employed women for each country in my sample. Though there is an observed negative relationship between the countries, this relationship appears to be stronger when comparing

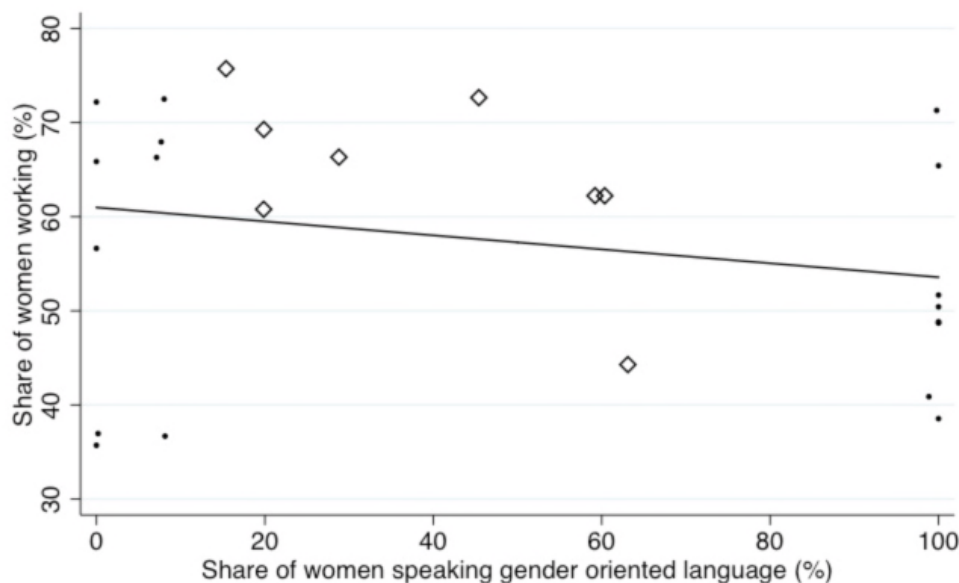


FIGURE 1. SHARE OF WOMEN WORKING AND SHARE OF WOMEN WITH NGII = 1 BY COUNTRY.

*Note:* Countries with greater than 10% variation in NGII are marked by the diamond marker. These include Cameroon, Namibia, Uganda, Benin, Nigeria, Togo, Kenya and Burkina Faso.

countries with high variation in NGII. This can be observed by considering the subset of 8 countries that have greater than 10% variation in NGII. Regressions with country fixed effects primarily rely on these countries with high within-country variation for identification. I therefore move forward by following Chen (2013) in testing the robustness of the results to estimation using different levels of NGII variation.

First, I test whether the relationship between NGII and employment is driven by variation in gender-oriented languages within the country. I run the main specification with the full set of individual controls and country fixed effects restricting the sample firstly, to those countries with less than 10% variation in NGII and secondly, to those countries with more than 10% variation. I present the results for this in [Table 9](#).

Though the estimated magnitude in regression (1) is still similar in magnitude to the main result, the reduction in signal-to-noise ratio due to lack of NGII variation means that the coefficient is statistically indistinguishable from zero. By contrast, results from regression (2) are strikingly similar to estimation using the whole sample: it is significant at 5% and the estimated coefficient is very close in magnitude to -0.0585 (from [Table 2](#)). As identification seems to be driven by countries with high NGII variation, in regression (3), I run the same specification

excluding Togo and Uganda, two countries with greater than 10% NGII variation. Though the standard error increases, the estimated coefficient is identical to that estimated with the full sample.

Table 9—: Regressions by NGII Variation

	(1) Employed	(2) Employed	(3) Employed	(4) Employed	(5) Employed	(6) Employed
NGII	-0.0501 (0.0758)	-0.0579** (0.0263)	-0.0585* (0.0311)	-0.0501 (0.0758)	-0.0579** (0.0262)	-0.0585** (0.0285)
NGII Share				-0.00125 (0.000804)	-0.005577** (0.00204)	-0.001155* (0.000619)
Sample used	<10% NGII Variation	>10% NGII Variation	All, excluding Togo and Uganda	<10% NGII Variation	>10% NGII Variation	All
Individual Controls	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES
Observations	152,445	114,440	239,949	152,445	114,440	266,885

*Note:* Standard errors clustered at level of the language are reported in parentheses. \*, \*\* and \*\*\* indicate the significance of the p-values at 10%, 5% and 1%.

In the next set of regressions, I run the main specification with controls for “NGII Share”, a measure for the number of people speaking the same NGII classification as the individual. Batibo (2005) finds that individuals speaking minority languages are disadvantaged in political and decision making environments. In regressions (4) to (6), I therefore consider the effect of “NGII Share” separately for those countries with <10% NGII variation, with >10% NGII Variation, and finally with the entire sample. As seen from regressions (4) to (6), addition of “NGII Share” leaves the coefficients virtually unchanged. This provides evidence that the main results are not driven by those who speak a minority NGII classification, or group effects of others speaking a language of the same degree of gender orientation.

## 11. Conclusion

My results suggest that women speaking gender oriented languages in Sub-Saharan Africa are on average 5 percentage points less likely to participate in the labour force. The effect of language also extends to gender segregation in the

labour force, in which I find of those who are working, women speaking gender oriented languages are 9.96 percentage points less likely to work in agriculture. In context with the existing literature, I find that the effect of language is lower in Sub-Saharan Africa than that of the global average. This is suggestive of language perhaps not having as strong an effect in areas with rich linguistic diversity.

On a broader scale, these results provide evidence for the importance of socio-psychological factors and gender identity in explaining differential labour market outcomes for women. My results suggest that language does not differentially affect a woman's material benefits to working. Moreover, the relationship between language and labour force participation does not seem to operate through gender social norms arising from contemporary religious beliefs and ancestral characteristics. Though I cannot fully account for gender social norms related to female labour supply, these results provide preliminary evidence of a direct effect of gender-oriented languages on the saliency of gender roles in women's minds.

The effect of language on labour force participation would have interesting policy implications, especially in countries with rich linguistic heterogeneity. From a linguistic standpoint, they hint at the need for the emphasis of less gender-oriented languages - for work, or for medium of education - in countries where there are already multiple official languages. In addition, the results are suggestive of the importance of implementing policies to counter the saliency of gender norms, such as positive discrimination in economic environments or media campaigns to blur the distinctions between male and female roles.

## REFERENCES

- Acemoglu, D., Johnson S., and Robinson. J. A.** 2001. *Reversal of fortune: Geography and institutions in the making of the modern world income distribution.*
- Acemoglu, D., Johnson S., and Robinson. J. A.** 2002. *The rise of Europe: Atlantic trade, institutional change and economic growth.*
- Aikhenvald, A. Y.** 2003. "Evidentiality in typological perspective." *Typological Studies in Language*, 54: 1–32.
- Akerlof, G. A., and R. E. Kranton.** 2000. "Economics and identity." *Quarterly Journal of Economics*, 715–753.
- Akerlof, G. A., and R. E. Kranton.** 2002. "Identity and schooling: Some lessons for the economics of education." *Journal of economic literature*, 40(4): 1167–1201.
- Akerlof, G. A., and R. E. Kranton.** 2005. "Identity and the Economics of Organizations." *The Journal of Economic Perspectives*, 19(1): 9–32.
- Alesina, A., and E. La Ferrara.** 2005. "Preferences for redistribution in the land of opportunities." *Journal of public Economics*, 89(5): 897–931.
- Alesina, A., Devleeschauwer A. Easterly W. Kurlat S., and R. Wacziarg.** 2003. "Fractionalization." *Journal of Economic growth*, 8(2): 155–194.
- Alesina, A., Giuliano P., and N. Nunn.** 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics*, 128(2).
- Altonji, J. G., Elder T. E., and C. R. Taber.** 2000. *Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools.*
- Anderson, K. L.** 1997. "Gender, status, and domestic violence: An integration of feminist and family violence approaches." *Journal of Marriage and the Family*, 655–669.
- Anderson, W. G.** 1897. *An Introductory Grammar of the Sena Language.* London:Northumberland Avenue, W.C.
- Angrist, J. D., and J.-S. Pischke.** 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.
- Baker, M. C.** 2003. *Lexical categories: Verbs, nouns and adjectives.* Vol. 102, Cambridge University Press. Vancouver.
- Bertrand, M.** 2011. "New perspectives on gender." *Handbook of labor economics*, 4: 1543–1590.
- Blau, F. D.** 1997. *Trends in the well-being of American women, 1970-1995.*
- Blau, F. D., and L. M. Kahn.** 2006. "The US gender pay gap in the 1990s: Slowing convergence." *Industrial labor relations review*, 60(1): 45–66.
- Bloom, D. E., Sachs J. D. Collier P., and C. Udry.** 1998. "Geography, demography, and economic growth in Africa." *Brookings papers on economic*

- activity*, 1998(2): 207–295.
- Bowern, C., and V. Lottridge.** 2002. *Ndebele*. Muenchen:Lincom Europa.
- Cagatay, N., and S. Ozler.** 1995. “Feminization of the labor force: The effects of long-term development and structural adjustment.” *World Development*, 23(11): 1883–1894.
- Camara, M.** 1999. *Parlons Malinke*. Paris:L’Harmattan.
- Cameron, A. C., and P. K. Trivedi.** 2005. *Microeconometrics: methods and applications*. Cambridge university press.
- Cameron, A. C., Gelbach J. B., and D. L. Miller.** 2008. “Bootstrap-based improvements for inference with clustered errors.” *The review of economics and statistics*, 90(3): 414–427.
- Cammenga, J.** 2002. *Phonology and Morphology of Ekegusii*. Gladback, B:Druckerei Hansen.
- Campante, F. R., and D. H. Yanagizawa-Drott.** 2013. *Does Religion Affect Economic Growth and Happiness? Evidence from Ramadan*.
- Charles, K. K., Guryan J. and Pan J.** 2009. “Sexism and womens labor market outcomes.” *Mimeo*.
- Chen, M. K.** 2012. “The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets.”
- Crazzolaro, P.** 1978. *A Study of the Pokon (Suk) Language: Grammar and Vocabulary*. Bologna:Editrice Missionaria Italiana.
- Dahl, O.** 2004. *The growth and maintenance of linguistic complexity*. Vol. 71, John Benjamins Publishing.
- Di Garbo, F.** 2013. “Evaluative morphology and noun classification: a cross-linguistic study of Africa.” Vol. 10, 114–146. SKASE.
- Dimmendaal, G. J.** 1983. *The Turkana Language*. Holland:Foris Publications.
- Dixon, R. MW.** 1986. “Noun classes and noun classification in typological perspective.” *Noun classes and categorization*, 105–112.
- Douglas, N.** 1991. “Institutions.” *Journal of Economic perspectives*, 5(1): 640–655.
- Dryer, M. S., and M. Haspelmath.** 1997. *Number of Genders, WALSL*.
- Duflo, E.** 2011a. *Women’s empowerment and economic development*.
- Duflo, E.** 2011b. *Women’s empowerment and economic development*.
- Easterly, W., and R. Levine.** 1997. “Africa’s growth tragedy: policies and ethnic divisions.” *The Quarterly Journal of Economics*, 1203–1250.
- Eckert, P., and S. McConnell-Ginet.** 2003. *Language and gender*. Cambridge University Press.
- Engerman, S. L., and K. L. Sokoloff.** 1997. “Factor endowments, institutions, and differential paths of growth among new world economies.” *How Latin America Fell Behind*, 260–304.

- Engerman, S. L., and K. L. Sokoloff.** 2002. *Factor endowments, inequality, and paths of development among new world economics.*
- Essien, E. O.** 1990. *A Grammar of the Ibibio Language.* Ibadan:University Press Limited.
- Farr, L., and F. Vella.** 2007. "The intergenerational transmission of gender role attitudes and its implications for female labor force participation."
- Faye, S.** 1982. *Morphologie Du Verbe Serere.* Centre de Linguistique Appliquee de Dakar.
- Fenske, J.** 2015. "African polygamy: Past and present." *Journal of Development Economics*, 117: 58–73.
- Fernndez, R.** 2007. *Culture as learning: The evolution of female labor force participation over a century.*
- Fortin, N. M.** 2005. "Gender role attitudes and the labour-market outcomes of women across OECD countries." *oxford review of Economic Policy*, 21(3): 416–438.
- Gay, V., Santacreu-Vasut E., and A. Shoham.** 2013. "The grammatical origins of gender roles."
- Girier, C.** 1996. *Parlons Soninke.* Paris:L'Harmattan.
- Guy, S.** 1978. *Labour force participation and development.* JSTOR.
- Haspelmeth, M., Dryer-M.S. Gil D., and B. Comrie.** 2005. "The world atlas of language structures."
- Heine, B.** 1982. "African noun class systems." *Apprehension: Das sprachliche Erfassen von Gegenstnden: I: Bereich und Ordnung der Phnomene*, 189–216.
- Hellinger, M., and H. Motschenbacher.** 2015. *Gender Across Languages.* Vol. Vol. 4, John Benjamins Publishing Company. Vancouver.
- Hopgood, C. R.** 1940. *Tonga Grammar.* London:Longmans, Green and Co. LTD.
- Jaumotte, F.** 2003. "Female labour force participation: past trends and main determinants in OECD countries."
- Kay, P., and W. Kempton.** 1984. "What is the SapirWhorf hypothesis?" *American anthropologist*, 86(1): 65–79.
- Klerman, D. M., Mahoney P. M. Spamann H., and M. I. Weinstein.** 2011. "Legal origin or colonial history?" *Journal of Legal Analysis.*
- Lai, V. T., and L. Boroditsky.** 2013. "The immediate and chronic influence of spatio-temporal metaphors on the mental representations of time in english, mandarin, and mandarin-english speakers." *Frontiers in psychology*, 4: 142.
- Last, J. T.** 1885. "Grammar of the Kamba language." *eastern equatorial Africa (London: Society for Promoting Christian Knowledge., 1885.*
- Lehrer, E.L.** 1995. "The effects of religion on the labor supply of married women." *Social science research*, 24(3): 281–301.

- Lehrer, E. L.** 2004. "Religion as a determinant of economic and demographic behavior in the United States." *Population and development review*, 30(4): 707–726.
- Lewis, M. P., Simons G.F., and C.D. Fennig**, ed. 2016. *Ethnologue: Languages of the World, Nineteenth edition*. Dallas, Texas: SIL International. Online version.
- Lopez de Silanes, F., La Porta R. Shleifer A., and R. Vishny.** 1998. "Law and finance." *Journal of political economy*, 106: 1113–1155.
- Macbrair, R. M.** 1842. *A grammar of the Mandingo language: with vocabularies*. Wesleyan-Methodist Missionary Society:London.
- Macmillan, R., and R. Gartner.** 1999. "When she brings home the bacon: Labor-force participation and the risk of spousal violence against women." *Journal of Marriage and the Family*, 947–958.
- Moulton, B.R.** 1986. "Random group effects and the precision of regression estimates." *Journal of Econometrics*, 32(3): 385–397.
- Nghiem, H.S., Nguyen H.T., R. Khanam, and L.B. Connelly.** 2015. "Does school type affect cognitive and non-cognitive development in children? Evidence from Australian primary schools." *Labour Economics*, 33: 55–65.
- Nunn, N.** 2014. "Historical development." *Handbook of Economic Growth*, 2: 347–402.
- Nunn, N., and L. Wantchekon.** 2009. *The slave trade and the origins of mistrust in Africa*.
- Ojiambo, P.** 2014. "University of Kansas, Lesson 9: Swahili Noun Classes."
- Olawsky, J. K.** 1999. *Aspects of Dagbani Grammar: with special emphasis on phonology and morphology*. Newcastle:Lincom Europa.
- Oster, E.** 2013. *Unobservable selection and coefficient stability: Theory and validation*.
- Price, T.** 1958. *The Elements of Nyanja*. Blantyre:Church of Scotland Mission. Vancouver.
- Saeed, J.** 1999. *Somali*. Vol. 10, John Benjamins Publishing.
- Sanderson, M.** 1922. *A Yao Grammar*. New York Toronto:The Macmillan Company.
- Soren, W., and E.W. Holman.** 2009. "Assessing temporal stability for linguistic typological features." *Mnchen: LINCOM Europa*.
- Tepowa, A.** 1904. "Notes on the (Nembe) Brass Language." *Journal of the Royal African Society*, 4(13): 117–133.
- Thompson, I.** 2015. *About World Languages*. Available at.
- Toweett, T.** 1979. *A study of Kalenjin Linguistics*. Nairobi:Kenya Literature Bureau.

- Tucker, A. N., and J. T. O. Mpaayei.** 1955. *A Maasai grammar, with vocabulary*. Green. Vancouver:Longmans.
- Vail, H. L.** 1971. "The noun classes of Tumbuka." *African studies*, 30(1): 35–60.
- Wald, B.** 1975. "Animate concord in Northeast Coastal Bantu: Its linguistic and social implications as a case of grammatical convergence." *Studies in African linguistics*, 6(3): 267–314.
- Welmers, W. E.** 1954. "Languages of West Africa." *Dietrich Westermann and MA Bryan. American Anthropologist*, 56(1): 148–150.
- Wilkie, M. B.** 1930. *Ga Grammar Notes and Exercises*. London:Humphrey Milford.
- Wilson, W. A. A.** 1962. "Temne, Landuma and the Baga languages." *Sierra Leone language review*, 1: 27–38.
- Zongo, B.** 2004. *Parlons Moore: Langue et culture des mossis Burkina-Faso*. Paris:L'Harmattan.

**APPENDIX A: Languages, Number of Noun Classes and Source**

<b>Language Name</b>	<b>Noun Classes</b>	<b>Source</b>	<b>Language Name</b>	<b>Noun Classes</b>	<b>Source</b>
Adja	0	Baker, 2003.	Moore	0	Zongo 2004
Akan	0	Hellinger and Motschenbacher 2015	Ndali	>5	Lewis et al. 2016
Amharic	2	Thompson 2015	Ndebele	>5	Bowern and Lottridge 2002
Arabic	2	Dryer and Haspelmath 1997	Nembe	3	Tepowa 1904
Bambara	0	Thompson 2015	Ngoni	>5	Dryer and Haspelmath 1997
Bariba	0	Lewis et al. 2016.	Nkhonde	>5	Lewis et al. 2016
Borana	2	Lewis et al. 2016.	Nyanja	>5	Price 1958
Chichewa	>5	Dryer and Haspelmath 1997	Oromo	2	Thompson 2015
Dagbani	>5	Olawsky 1999	Pokot	3	Crazzolaro 1978
Diola	>5	Dryer and Haspelmath 1997	Poular	>5	Dryer and Haspelmath 1997
Embu	>5	Wald 1975	Sena	>5	Anderson 1897
English	3	Dryer and Haspelmath 1997	Senufo	>5	Welmers 1954
Ewe	0	Dryer and Haspelmath 1997	Serer	0	Faye 1982
Fon	0	Baker 2003	Shona	>5	Dryer and Haspelmath 1997
French	2	Dryer and Haspelmath 1997	Somali	2	Thompson 2015
Fulani	>5	Dryer and Haspelmath 1997	Soninke	0	Grier 1996
Ga	>5	Wilkie 1930	Sukwa	>5	Lewis et al. 2016
Hausa	2	Dryer and Haspelmath 1997	Swahili	>5	Dryer and Haspelmath 1997
Ibibio	0	Essien 1990	Temne	>5	Wilson 1962
Igbo	0	Dryer and Haspelmath 1997	Tigrigna	2	Thompson 2015
Kalenjin	>5	Toweett 1979	Tiv	>5	Welmers 1974
Kamba	0	Last 1885	Tonga	>5	Hopgood 1940
Kikuyu	>5	Thompson 2015	Tumbuka	>5	Vail 1971
Kisii	>5	Cammenga 2002	Turkana	2	Dimmendaal 1983
Krio	0	Thompson 2015	Wolof	>5	Thompson 2015
Luo	0	Thompson 2015	Yao	>5	Sanderson1922
Maasai	3	Tucker and Mpaayei 1955.	Swahili	>5	Dryer and Haspelmath 1997
Malinka	2	Camara 1999	Temne	>5	Wilson 1962
Mandingue	0	Macbrair 1842	Tigrigna	2	Thompson 2015
Mang'anja	>5	Dryer and Haspelmath 1997	Tiv	>5	Welmers 1974
Marka	0	Galtier 1971	Tonga	>5	Hopgood 1940
Mende	0	Thompson 2015	Tumbuka	>5	Vail 1971
Meru	>5	Thompson 2016	Turkana	2	Dimmendaal 1983