Gendered Language

Pamela Jakiela  
CGD, UMD, BREAD, & IZA

Owen Ozier  
World Bank, BREAD, & IZA

October 2018
Motivation

Language structures thought:

“Languages differ not only in how they build their sentences but also in how they break down nature to secure elements to put in those sentences.”

– Benjamin Lee Whorf (1941)
Motivation

Language structures thought:

“Languages differ not only in how they build their sentences but also in how they break down nature to secure elements to put in those sentences.”

– Benjamin Lee Whorf (1941)

Sapir-Whorf Hypothesis: linguistic determinism

- Our native language limits the scope of our thinking
- Example [now debunked]: the Inuit have 7 different words for snow
Motivation

Language structures thought:

“Languages differ not only in how they build their sentences but also in how they break down nature to secure elements to put in those sentences.”

– Benjamin Lee Whorf (1941)

Sapir-Whorf Hypothesis: linguistic determinism

• Our native language limits the scope of our thinking

• Example [now debunked]: the Inuit have 7 different words for snow

Nonetheless, there is mounting evidence that the languages we speak influence our thoughts and actions in subtle, subconscious ways
Motivation: Language Structures Thought

**Example:** agentive language impacts perceptions of responsibility

Agentive language is the norm in English:

“Non-agentive language sounds evasive in English, the province of guilt-shirking children and politicians.”
Motivation: Language Structures Thought

Example: agentive language impacts perceptions of responsibility

Agentive language is the norm in English:

“Non-agentive language sounds evasive in English, the province of guilt-shirking children and politicians.”

Agentive language less common in other languages (e.g. Spanish)
Fausey and Boroditsky (2011) show Spanish, English monolinguals videos depicting “intentional” and “accidental” versions of the same event.

<table>
<thead>
<tr>
<th>Action</th>
<th>Intentional</th>
<th>Accidental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knocks box</td>
<td>Faces table, knocks box off table</td>
<td>Knocks box off table while gesturing</td>
</tr>
<tr>
<td>Breaks pencil</td>
<td>Sits at table, breaks pencil in half</td>
<td>Breaks pencil in half while writing</td>
</tr>
</tbody>
</table>
Fausey and Boroditsky (2011) show Spanish, English monolinguals videos depicting “intentional” and “accidental” versions of the same event

<table>
<thead>
<tr>
<th>Action</th>
<th>Intentional</th>
<th>Accidental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knocks box</td>
<td>Faces table, knocks box off table</td>
<td>Knocks box off table while gesturing</td>
</tr>
<tr>
<td>Breaks pencil</td>
<td>Sits at table, breaks pencil in half</td>
<td>Breaks pencil in half while writing</td>
</tr>
</tbody>
</table>

**Experiment 1:** subjects describe what happened

- Spanish-speakers less likely to use agentive language to describe accidental events, no differences observed for intentional acts
Motivation: Language Structures Thought

Fausey and Boroditsky (2011) show Spanish, English monolinguals videos depicting “intentional” and “accidental” versions of the same event

<table>
<thead>
<tr>
<th>Action</th>
<th>Intentional</th>
<th>Accidental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knocks box</td>
<td>Faces table, knocks box off table</td>
<td>Knocks box off table while gesturing</td>
</tr>
<tr>
<td>Breaks pencil</td>
<td>Sits at table, breaks pencil in half</td>
<td>Breaks pencil in half while writing</td>
</tr>
</tbody>
</table>

Experiment 1: subjects describe what happened

- Spanish-speakers less likely to use agentive language to describe accidental events, no differences observed for intentional acts

Experiment 2: (other) subjects try to remember who did what

- Both groups equally likely to remember intentional actors; Spanish-speakers less likely to remember who caused accidents
Motivation: Language Structures Thought

“Languages are far from impartial ‘containers’ for the packaging of underlying thoughts, but rather are active players in the construction of those thoughts.”

– Ogunnaike et al. (2010)

Language shapes our thoughts and actions in subtle ways:

- Using agentive language makes actors, responsibility more salient (Fausey et al. 2010, Fausey and Boroditsky 2011)

- Language of instructions, stimuli impacts implicit prejudices of bilinguals (Danzinger and Ward 2009, Ogunnaike et al. 2010)

- Speakers of languages that treat the future as a separate tense save less than those that treat the future like the present (Chen 2013)
Motivation: Gender Norms

Percent agreeing: **when jobs are scarce, men should have more of a right to a job than women**

Percent agreeing: **when a woman works, the children suffer**

Motivation: Gender and Language

Linguistic gender distinctions:

• Pronominal distinctions between men and women

• Nominal classification systems (grammatical gender)
Motivation: Gender and Language

Linguistic gender distinctions:

- Pronominal distinctions between men and women
- Nominal classification systems (grammatical gender)

Do linguistic gender distinctions impact gender norms?

*Grammatical gender creates “a habitual consciousness of two sex classes as a standard classificatory fact in our thought-world.”*

– Benjamin Lee Whorf (1936)

Builds on arguments advanced by Durkheim and Mauss (1903)
Motivation: Gender and Language

Suggestive evidence of a link between grammar and gender norms:

- Givati and Troiano (2012) show that countries with gendered pronouns have shorter government-mandated maternity leaves
- Perez and Tavits (2018) show that grammatical gender impacts gender attitudes among Estonian/Russian bilinguals
- Hicks *et al.* (2015) use WALS data to look at US immigrants
Our Contribution

1. Characterize the grammatical gender structure of 4,334 languages which together account for 99 percent of the world’s population
   - India: 6 languages coded in WALS, we code 281
   - Kenya: 3 languages coded in WALS, we code (all) 51

2. Estimate the proportion of each country’s population whose native language uses a grammatical gender system to classify nouns
   - Estimate the cross-country relationship between grammatical gender and women’s labor force participation and educational attainment, and the relationship with gender attitudes among men and women

3. Use individual-level data from countries where both gender and non-gender languages are indigenous and widely spoken
   - Replicate cross-country results within countries
Outline of the Talk

1. What is grammatical gender?
2. Identifying gender languages
3. Cross-country analysis
   3.1 Labor force participation
   3.2 Educational attainment
   3.3 Gender attitudes
4. Within-country analysis
   4.1 Labor force participation
   4.2 Educational attainment
5. Discussion, policy implications, and conclusion
Non-Preview of Main Results
Grammatical Gender
Motivation: Gender and Language

Languages differ in their treatment of gender:

- Pronominal distinctions between men and women
- Nominal classification systems (grammatical gender)
Motivation: Gender and Language

Languages differ in their treatment of gender:

- Pronominal distinctions between men and women
- Nominal classification systems (grammatical gender)

**Example:** Swahili does not make pronominal gender distinctions

\[
\begin{align*}
\text{she} & \text{ goes to school} \\
\text{he} & \text{ goes to school}
\end{align*}
\]

\[\text{[yeye] anaenda shuleni}\]
Motivation: Gender and Language

Languages differ in their treatment of gender:

- Pronominal distinctions between men and women
- Nominal classification systems (grammatical gender)

Example: Swahili does not make pronominal gender distinctions

\[
\begin{align*}
\text{she goes to school} & \quad [\text{yeye}] \text{ anaenda shuleni} \\
\text{he goes to school} & \quad [\text{yeye}] \text{ anaenda shuleni}
\end{align*}
\]

There are different words for males and females (e.g. “boy” vs. “girl”), but genders are treated identically from a grammatical perspective.
Motivation: Gender and Language

Example: Spanish uses different pronouns for males and females

she goes to school = ella va a la escuela
he goes to school = él va a la escuela
Motivation: Gender and Language

Example: Spanish uses different pronouns for males and females

she goes to school = ella va a la escuela
he goes to school = él va a la escuela

Spanish uses a system of **grammatical gender** to classify nouns

- All Spanish nouns are either masculine or feminine
- Grammatical gender determines agreement (e.g. with adjectives)
Nominal classification

Most languages have a system for categorizing nouns (Aikhenvald 2003)

- Many languages partition nouns into **noun classes** or genders
Nominal classification

Most languages have a system for categorizing nouns (Aikhenvald 2003)

- Many languages partition nouns into **noun classes** or genders

Elements of a noun class often often often share morphological properties:

- **Spanish:**
  - Masculine words end in O
  - Feminine words end in A

- **Swahili:** class prefixes are used as class names
  - small items belong in ki-/vi- class, humans in m-/wa- class
Nominal classification

Most languages have a system for categorizing nouns (Aikhenvald 2003)

- Many languages partition nouns into **noun classes** or genders

Elements of a noun class often often often share morphological properties:

- **Spanish:**
  - Masculine words end in O
  - Feminine words end in A

- **Swahili:** class prefixes are used as class names
  - small items belong in ki-/vi- class, humans in m-/wa- class

Typical noun class system = semantic core + many exceptions
Nominal classification

Noun classes are defined by agreement — eg. nouns with adjectives

Example: Swahili has nine distinct noun classes, each characterized by a set of prefixes for verbs, adjectives, demonstratives, possessives, etc.

- these new chairs = viti vipya hivi
- these new teachers = walimu wapya hawa

Example: agreement depends on gender (masc./fem.) in Spanish

- the white shirt = la camisa blanca
- the white hat = el sombrero blanco
A **grammatical gender** system is a system of noun classification that:

- Includes masculine and feminine as two of the classes
- Characterizes (some) inanimate objects as masculine or feminine
  - English is not a gender language* (though it uses gender pronouns)
A grammatical gender system is a system of noun classification that:

- Includes masculine and feminine as two of the classes
- Characterizes (some) inanimate objects as masculine or feminine
  - English is not a gender language* (though it uses gender pronouns)

Languages that use grammatical gender — a.k.a. gender languages — differ in grammatical gender intensity along several dimensions

- Do the masculine and feminine classes partition the noun space?
  - Many languages have a neuter class (eg. German, Russian)
- How many parts of speech must change to reflect agreement?
  - Example: verbs agree with gender in Russian, but not in Spanish
Does Grammatical Gender Matter?

Conventional wisdom is that grammatical gender is arbitrary:

“In German, a young lady has no sex, while a turnip has.”

– Mark Twain
Does Grammatical Gender Matter?

Conventional wisdom is that grammatical gender is arbitrary:

“In German, a young lady has no sex, while a turnip has.”

– Mark Twain

Some linguists have questioned this assumption (cf. Lakoff 1987), arguing that gender categories have a certain cultural intelligibility

• In Dyirbal, women are grouped with fire and “dangerous things”

• In Ket, one linguist suggested that certain small mammals are feminine “because they are of no importance to the Kets”

• Assignment of inanimate objects to grammatical gender categories often reflects stereotypes about male vs. female body types
Does Grammatical Gender Matter?

Native German speakers said: Native Spanish speakers said:
hard golden
heavy intricate
jagged little
metal lovely
serrated shiny
der Schlüssel la llave
(masculine) (feminine)

Source: Boroditsky et al. (2002)

Does Grammatical Gender Matter?

Native German speakers said:
- hard
- heavy
- jagged
- metal
- serrated

Native Spanish speakers said:
- golden
- intricate
- little
- lovely
- shiny

Source: Boroditsky et al. (2002)

Jakiela and Ozier (2018)
Gendered Language, Slide 19
Does Grammatical Gender Matter?

Native **German** speakers said:
- hard
- heavy
- jagged
- metal
- serrated

Native **Spanish** speakers said:
- golden
- intricate
- little
- lovely
- shiny

**der Schlüssel**
(masculine)

**la llave**
(feminine)

Source: Boroditsky *et al.* (2002)

Does Grammatical Gender Matter?

Native **German** speakers said:  

- beautiful  
- elegant  
- fragile  
- peaceful  
- pretty

Native **Spanish** speakers said:

- big  
- dangerous  
- long  
- strong  
- sturdy

Source: Boroditsky et al. (2002)  

Jakiela and Ozier (2018)  

Gendered Language, Slide 20
Does Grammatical Gender Matter?

Native German speakers said:
- beautiful
- elegant
- fragile
- peaceful
- pretty

Native Spanish speakers said:
- big
- dangerous
- long
- strong
- sturdy

**die Brücke**  
*(feminine)*

**el puente**  
*(masculine)*

Source: Boroditsky et al. (2002)
Does Grammatical Gender Matter?

Evidence that grammatical gender matters:

- Santacreu-Vasut *et al.* (2013): political quotas for women are more common in countries where the national language is non-gender

- Hicks *et al.* (2015): immigrants are more likely to divide household tasks along gender lines if they grew up speaking a gender language

- Perez and Tavits (2018): Estonian/Russian bilinguals show greater support for gender equality when (randomly) interviewed in Estonian
Evidence that grammatical gender matters:

- Santacreu-Vasut *et al.* (2013): political quotas for women are more common in countries where the national language is non-gender.

- Hicks *et al.* (2015): immigrants are more likely to divide household tasks along gender lines if they grew up speaking a gender language.

- Perez and Tavits (2018): Estonian/Russian bilinguals show greater support for gender equality when (randomly) interviewed in Estonian.

Studies suggest grammatical gender associated with women’s exclusion from public life, labor market, etc.; specialization in domestic sphere.

- Existing work hampered by data limitations.
Identifying Gender Languages
The Ethnologue is the most comprehensive database of languages

- Includes over 7,000; 6,190 of them living oral native languages
In many (LIC/LMIC) countries, the most widely spoken native language accounts for a small fraction of the population (e.g. 0.18 in Nigeria)
Classifying Gender Structures

We compile data on grammatical structures from a range of sources:

- World Atlas of Language Structures
Classifying Gender Structures

We compile data on grammatical structures from a range of sources:

- World Atlas of Language Structures
- *Linguistic Survey of India*
  - Compiled by George A. Grierson between 1891 and 1928
- George L. Campbell’s *Compendium of the World’s Languages*
- Language-specific data sources:
  - Grammatical monographs
  - Language textbooks and online learning materials
  - Academic work by (modern) linguists
  - Interviews with native speakers and translators
Classifying Gender Structures

For each language, we attempt to code two variables:

- A indicator for using any system of grammatical gender
- A indicator for using a dichotomous system of grammatical gender
  - All nouns must be either masculine or feminine
Classifying Gender Structures

For each language, we attempt to code two variables:

- A indicator for using any system of grammatical gender
- A indicator for using a dichotomous system of grammatical gender
  - All nouns must be either masculine or feminine

We do not attempt to determine:

- The number of genders/classes, if there are more than two
- The intensity of the agreement system (i.e. what must agree)
- The presence of gendered personal pronouns (for humans)
Classifying Gender Structures

Languages positively identified as **gender** languages in two ways:

1. Explicit statement about grammatical gender structure

**Serbian**: “Three grammatical genders (masculine, feminine, and neuter) and two numbers (singular and plural) are also distinguished.”

**Tigrinya**: “Tigrinya nouns are either masculine or feminine and are inflected for number. Gender is not marked on the noun, but on nominal dependents like articles and adjectives. Verbs agree with their subjects and objects in person, number, and gender.”
Classifying Gender Structures

Languages positively identified as gender languages in two ways:

1. Explicit statement about grammatical gender structure

**Serbian:** “Three grammatical genders (masculine, feminine, and neuter) and two numbers (singular and plural) are also distinguished.”

**Tigrinya:** “Tigrinya nouns are either masculine or feminine and are inflected for number. Gender is not marked on the noun, but on nominal dependents like articles and adjectives. Verbs agree with their subjects and objects in person, number, and gender.”

2. A textbook or language-specific grammar indicates that:
   - There are masculine and feminine noun classes (genders), at least one of which includes nouns other than male/female animates
   - Adjectives or another part of speech must agree in gender

Classifying Gender Structures

Languages identified as **non-gender** languages in the same ways:

1. Explicit statement about grammatical gender structure

   **Gamo:** “The use of gender is governed by non-linguistic factors — i.e. by the actual sex of the referent.”

   **Maithili:** “Modern Maithili, however, has no grammatical gender. In other words, in modern Maithili, distinctions of gender are determined soley by the sex of the animate noun.”

   **Nuosu:** “There is no grammatical gender, and such words as do not denote animate beings have no gender at all.”
Classifying Gender Structures

Languages identified as non-gender languages in the same ways:

1. Explicit statement about grammatical gender structure

   **Gamo:** “The use of gender is governed by non-linguistic factors — i.e. by the actual sex of the referent.”

   **Maithili:** “Modern Maithili, however, has no grammatical gender. In other words, in modern Maithili, distinctions of gender are determined soley by the sex of the animate noun.”

   **Nuosu:** “There is no grammatical gender, and such words as do not denote animate beings have no gender at all.”

2. A textbook or language-specific grammar describes nouns or nominals without mentioning any noun class system, or describes a system of classes that do not include either masculine or feminine
We identify the grammatical structure of 4,334 of 6,190 languages

- All but four of the languages with more than one million speakers
- Verify gender structure w/ two sources whenever possible
Classifying Gender Structures

We classify more than 95 percent of population in all but eight countries.
The Distribution of Gender Languages

Native speakers of gender languages: 38 percent of world’s population

→ [Comparison with WALS]
Comparing country-level WALS data to full data

Two measures of the fraction of a country speaking a gender language as their native language

Cross-Country Analysis
Cross-Country Analysis: Data

1. Labor force participation
   - World Development Indicators
   - Available for 177 countries
Cross-Country Analysis: Data

1. Labor force participation
   - World Development Indicators
   - Available for 177 countries

2. Educational attainment (primary and secondary school completion)
   - Barro-Lee Educational Attainment Data
   - Available for 142 countries
Cross-Country Analysis: Data

1. Labor force participation
   - World Development Indicators
   - Available for 177 countries

2. Educational attainment (primary and secondary school completion)
   - Barro-Lee Educational Attainment Data
   - Available for 142 countries

3. Gender attitudes
   - World Values Survey, Round 6
   - Available for 56 countries
Cross-Country Analysis: Empirical Specifications

We estimate OLS regressions of the form:

\[ Y_c = \alpha + \beta Gender_c + \delta_{continent} + \lambda X_c + \varepsilon_c \]

where:

- \( Gender_c \) is the proportion of population speaking gender language
- \( \delta_{continent} \) is a vector of continent fixed effects
- \( X_c \) is a vector of country-level geographic controls:
  - Average rainfall, average temperature, proportion tropical, indicator for being landlocked, suitability for the plough
- \( \varepsilon_c \) is a mean-zero error term
1. Measurement error in country-level prevalence of gender languages
   ▶ Bounding exercise following Imbens and Manski (2004)
Cross-Country Analysis: Robust Inference

1. Measurement error in country-level prevalence of gender languages
   ▶ Bounding exercise following Imbens and Manski (2004)

2. Non-independence of languages with families
   ▶ Permutation test based on structure of the language tree
Cross-Country Analysis: Assessing Causality

1. Examine within-country gender differences, where applicable
   - Applies to LFP and education, not gender attitudes
Cross-Country Analysis: Assessing Causality

1. Examine within-country gender differences, where applicable
   ▶ Applies to LFP and education, not gender attitudes

2. Examine coefficient stability, robustness to observable controls
   ▶ Follow Altonji et al. (2005), Oster (forthcoming)
Cross-Country Analysis: Assessing Causality

1. Examine within-country gender differences, where applicable
   ▶ Applies to LFP and education, not gender attitudes
2. Examine coefficient stability, robustness to observable controls
   ▶ Follow Altonji et al. (2005), Oster (forthcoming)
3. Replicate cross-country results using within-country variation
Cross-Country Analysis: Female LFP

Proportion gender < 0.1

0.1 < proportion gender < 0.9

Proportion gender > 0.9

Jakiela and Ozier (2018) Gendered Language, Slide 41
## Cross-Country Analysis: Female LFP

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>( \text{LFP}_f )</th>
<th>( \text{LFP}_f - \text{LFP}_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification:</td>
<td>OLS (1)</td>
<td>OLS (2)</td>
</tr>
<tr>
<td></td>
<td>OLS (3)</td>
<td>OLS (4)</td>
</tr>
<tr>
<td>Proportion gender</td>
<td>-13.83 (2.80)</td>
<td>-11.61 (2.47)</td>
</tr>
<tr>
<td></td>
<td>([p &lt; 0.001])</td>
<td>([p &lt; 0.001])</td>
</tr>
<tr>
<td></td>
<td>-11.92 (3.34)</td>
<td>-14.66 (3.25)</td>
</tr>
<tr>
<td></td>
<td>([p &lt; 0.001])</td>
<td>([p &lt; 0.001])</td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country-Level Geography Controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>178</td>
<td>178</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.15</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Robust standard errors are clustered by the most widely spoken language in all specifications; they are reported in parentheses. P-values are reported in square brackets. \( \text{LFP}_f \) is the percentage of women in the labor force, measured in 2011. \( \text{LFP}_f - \text{LFP}_m \) is the gender difference in labor force participation — i.e. the difference between female and male labor force participation, again measured in 2011. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina et al. (2013) measure of suitability for the plough.
Cross-Country Analysis: Female LFP

Estimated coefficients are economically significant:

- Grammatical gender could fully explain the disparity in female labor force participation between Jamaica and the Dominican Republic

- Grammatical gender keeps 125 million women out of work force
Cross-Country Analysis: Female LFP

Robustness checks:

- Marginal impact of stronger grammatical gender systems
- Including “bad” controls
- Omitting major world languages
Primary education by continent

Jakiela and Ozier (2018)  Gendered Language, Slide 45
Cross-Country Analysis: Educational Attainment

→ Secondary education by continent
## Cross-Country Analysis: Educational Attainment

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>PRI&lt;sub&gt;f&lt;/sub&gt;</th>
<th>PRI&lt;sub&gt;f&lt;/sub&gt; - PRI&lt;sub&gt;m&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specification:</strong></td>
<td>OLS (1)</td>
<td>OLS (2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Proportion gender</td>
<td>14.79</td>
<td>-6.71</td>
</tr>
<tr>
<td></td>
<td>(5.83)</td>
<td>(4.40)</td>
</tr>
<tr>
<td></td>
<td>[0.013]</td>
<td>[0.130]</td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country-Level Geography Controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>142</td>
<td>142</td>
</tr>
<tr>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.06</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Robust standard errors are clustered by the most widely spoken language in all specifications; they are reported in parentheses. P-values are reported in square brackets. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina *et al.* (2013) measure of suitability for the plough.
Cross-Country Analysis: Educational Attainment

<table>
<thead>
<tr>
<th>Specification</th>
<th>SEC&lt;sub&gt;f&lt;/sub&gt;</th>
<th>SEC&lt;sub&gt;f&lt;/sub&gt; - SEC&lt;sub&gt;m&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
<td>OLS (2)</td>
</tr>
<tr>
<td>Proportion gender</td>
<td>14.52 (5.77)</td>
<td>0.43 (3.70)</td>
</tr>
<tr>
<td></td>
<td>[0.013]</td>
<td>[0.907]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.48 (1.93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.802]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.86 (2.35)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.716]</td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country-Level Geography Controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>142</td>
<td>142</td>
</tr>
<tr>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.06</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Robust standard errors are clustered by the most widely spoken language in all specifications; they are reported in parentheses. P-values are reported in square brackets. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina et al. (2013) measure of suitability for the plough.
World Values Survey includes 8 questions on gender attitudes:

- When a mother works for pay, the children suffer [1]
- When jobs are scarce, men should have more right to a job than women [1]
- On the whole, men make better political leaders than women do [1]
- On the whole, men make better business executives than women do [1]
- Being a housewife is just as fulfilling as working for pay [1]
- If a woman earns more money than her husband, it’s almost certain to cause problems [1]
- A university education is more important for a boy than for a girl [1]
- Having a job is the best way for a woman to be an independent person [0]
Cross-Country Analysis: Gender Attitudes

- Men make better political leaders: p = 0.006
- Men have more right to a scarce job: p = 0.012
- Men make better business executives: p = 0.005
- When a mother works, the children suffer: p = 0.009
- Being a housewife as fulfilling as paid work: p = 0.042
- If a wife earns more, it causes problems: p = 0.081
- University is more important for boys: p = 0.005
- Having a job not best way to be independent: p = 0.685

Jakiela and Ozier (2018) Gendered Language, Slide 50
Cross-Country Analysis: Gender Attitudes

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Gender Attitude Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification:</td>
<td>OLS (1)</td>
</tr>
<tr>
<td></td>
<td>OLS (2)</td>
</tr>
<tr>
<td>Proportion gender</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td></td>
<td>[0.576]</td>
</tr>
<tr>
<td></td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td>[0.002]</td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>No</td>
</tr>
<tr>
<td>Country-Level Geography Controls</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>56</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>0.78</td>
</tr>
</tbody>
</table>

Robust standard errors clustered by most widely spoken language in all specifications. The Gender Attitude Index is the first principal component of responses to the eight questions on gender attitudes included in the World Values Survey. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina et al. (2013) measure of suitability for the plough.
Cross-Country Analysis: Gender Attitudes

Attitudes among Women:

- Men make better political leaders
- Men have more right to a scarce job
- Men make better business executives
- When a mother works, the children suffer
- Being a housewife as fulfilling as paid work
- If a wife earns more, it causes problems
- University is more important for boys
- Having a job not best way to be independent

Attitudes among Men:

- Men make better political leaders
- Men have more right to a scarce job
- Men make better business executives
- When a mother works, the children suffer
- Being a housewife as fulfilling as paid work
- If a wife earns more, it causes problems
- University is more important for boys
- Having a job not best way to be independent

## Cross-Country Analysis: Gender Attitudes

<table>
<thead>
<tr>
<th>Sample:</th>
<th>Attitude Index: Women</th>
<th></th>
<th>Attitude Index: Men</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification:</td>
<td>OLS (1)</td>
<td>OLS (2)</td>
<td>OLS (3)</td>
<td>OLS (4)</td>
</tr>
<tr>
<td>Proportion gender</td>
<td>-0.02 (0.05)</td>
<td>-0.10 (0.04)</td>
<td>-0.04 (0.06)</td>
<td>-0.14 (0.04)</td>
</tr>
<tr>
<td>[0.714]</td>
<td>[0.012]</td>
<td>[0.508]</td>
<td>$p &lt; 0.001$</td>
<td></td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Geography Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.00</td>
<td>0.73</td>
<td>0.02</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Robust standard errors clustered by most widely spoken language in all specifications. The **Gender Attitude Index** is the first principal component of responses to the eight questions on gender attitudes included in the World Values Survey. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina et al. (2013) measure of suitability for the plough.
Cross-Country Analysis: Measurement Error

The problem: RHS variable is an interval for 85 of 193 countries

- Analysis thus far assumes missingness is ignorable
- Measurement error is not classical, could bias estimates
Cross-Country Analysis: Measurement Error

The problem: RHS variable is an interval for 85 of 193 countries

• Analysis thus far assumes missingness is ignorable
• Measurement error is not classical, could bias estimates

Our approach: calculate bounds following Imbens and Manski (2004)

1. Identify highest and lowest coefficient estimates numerically
2. Calculate associated naïve confidence intervals, take the union
3. Symmetrically tighten the confidence interval for correct coverage
Cross-Country Analysis: Measurement Error

Full data vs WALS-only data

Women's Attitudes

Men's Attitudes

Attitude Index

PRIfemale - PRImale

LFPfemale - LFPmale

PRIfemale - PRImale

Attitude Index

Men's Attitudes

Women's Attitudes

95 percent confidence interval

Naive OLS CI

Imbens-Manski CI

→ [Manski table]

Jakiela and Ozier (2018)  Gendered Language, Slide 56
The problem: languages are not independent (Roberts et al. 2015)

- Useful variation in grammatical structure within and between families
- Intuitively, this is a clustering problem, but countries not nested
Cross-Country Analysis: Independence

The problem: languages are not independent (Roberts et al. 2015)
- Useful variation in grammatical structure within and between families
- Intuitively, this is a clustering problem, but countries not nested

Our approach: permutation tests based on the language tree

1. Assign languages to largest possible homogeneous clusters
2. Randomly permute treatment (grammatical gender) across clusters
3. Replicate cross-country analysis for each hypothetical treatment
   \[\Rightarrow\] Allows us to calculate permutation-test p-values
Cross-Country Analysis: Permutation Tests

Dravidian
- Southern
  - Tulu
  - Koraga
- Tamil-Kannada
  - Kannada
  - Tamil-Kodagu
  - Tamil-Malayalam
- Tamil

Kodava
- Malayalam
- Taniya
- Rayula
- Irula
- Tamil
- Verukula

Adilabad Gondi
- Gondi
- Aheri Gondi
- Northern Gondi
- Konda-Dora
- Koya
- Kuvi
- Mukha-Dora

Northern Gondi
- Gondi-Kui
- Duruwa
- Pottangi Ollar Gadda

Northern
- Sauria Paharia
- Kurux
- Kumarbhag Paharia

Central
- Gondi-Kui
- Kolami-Naiki
- Duruwa
- Pottangi Ollar Gadda

South-Central
- Telugu

Northern
- Sauria Paharia
- Kurux
- Kumarbhag Paharia

Dravidian
- Tulu
- Koraga
- Korra Koraga

Cross-Country Analysis: Permutation Tests

Dravidian
- Southern
  - Tulu
  - Koraga
  - Korra
- Tamil-Kannada
  - Kannada
  - Tamil-Kodagu
  - Tamil-Malayalam
  - Tamil
  - Malayalam
- Konda-Koi
  - Konda-Dora
  - Koya
  - Kui
  - Mukha-Dora
- Gondi
  - Adilabad Gondi
  - Aheri Gondi
  - Northern Gondi
- Parji-Gadaba
  - Pottangi Ollar Gadaba
- Gondi-Kui

Central
- Telugu
- Konda-Koi

North-Central
- Dravidian

General
Cross-Country Analysis: Permutation Tests

Female LFP:

Gender Difference in LFP:
### Cross-Country Analysis: Permutation Tests

<table>
<thead>
<tr>
<th></th>
<th>Naïve OLS p-values</th>
<th>Permutation-based p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female labor force participation</td>
<td>0.00050</td>
<td>0.01520</td>
</tr>
<tr>
<td>Gender difference in labor force participation</td>
<td>0.00001</td>
<td>0.00810</td>
</tr>
<tr>
<td>Female primary school completion</td>
<td>0.13012</td>
<td>0.16920</td>
</tr>
<tr>
<td>Gender difference in primary school completion</td>
<td>0.08773</td>
<td>0.08820</td>
</tr>
<tr>
<td>Female secondary school completion</td>
<td>0.90692</td>
<td>0.92410</td>
</tr>
<tr>
<td>Gender difference in secondary school completion</td>
<td>0.71638</td>
<td>0.73140</td>
</tr>
<tr>
<td>Gender attitudes index</td>
<td>0.00225</td>
<td>0.05030</td>
</tr>
<tr>
<td>Gender attitudes index among women</td>
<td>0.01223</td>
<td>0.09620</td>
</tr>
<tr>
<td>Gender attitudes index among men</td>
<td>0.00063</td>
<td>0.03040</td>
</tr>
</tbody>
</table>

P-values estimated using 10,000 permutations. For each outcome, the naïve p-value comes from the associated regression in a previous table. The permutation-based p-value is the fraction of permutations in which the magnitude of the estimated coefficient (from a hypothetical permutation of the gender indicator that respects the cluster structure of the language tree) exceeds the magnitude of the estimated coefficient in the true (non-permuted) data set.
Cross-Country Analysis: Coefficient Stability

Altonji et al. (2005) and Oster (2017) propose using robustness to observable controls to assess the magnitude of omitted variable bias

- Bias from unobservables is proportional to coefficient movements
- Coefficient movements must be scaled by changes in $R^2$
Altonji et al. (2005) and Oster (2017) propose using robustness to observable controls to assess the magnitude of omitted variable bias

- Bias from unobservables is proportional to coefficient movements
- Coefficient movements must be scaled by changes in $R^2$

Consider a true model:

$$Y = \alpha + \beta X + \eta W_{\text{observable}} + \gamma W_{\text{unobservable}} + \varepsilon$$
Cross-Country Analysis: Coefficient Stability

Altonji et al. (2005) and Oster (2017) propose using robustness to observable controls to assess the magnitude of omitted variable bias

• Bias from unobservables is proportional to coefficient movements

• Coefficient movements must be scaled by changes in $R^2$

Consider a true model:

$$Y = \alpha + \beta X + \eta W_{\text{observable}} + \gamma W_{\text{unobservable}} + \varepsilon$$

Data on $Y$, $X$, and $W_{\text{observable}}$ tells us:

• How much does $\hat{\beta}$ change when $W_{\text{observable}}$ is included?

• How much of the residual variation in $Y$ is explained by $W_{\text{observable}}$?
In this framework, $\delta$ is a proportional selection coefficient:

$\delta$ denotes ratio of (i) covariance between treatment and unobserved controls to (ii) covariance between treatment and observed controls.
In this framework, $\delta$ is a **proportional selection coefficient**:

$\delta$ denotes ratio of (i) covariance between treatment and unobserved controls to (ii) covariance between treatment and observed controls

Regression results with and without controls allow us to calculate:

- True causal $\beta^*$ under the assumption that $\delta = 1$
- Value of $\delta^*$ that would be required for omitted variable bias from unobservables to fully explain observed association between $X$ and $Y$
  - Altonji *et al.* (2005) suggest results are robust if $\delta > 1$
## Cross-Country Analysis: Coefficient Stability

### OLS Coefficients

<table>
<thead>
<tr>
<th></th>
<th>( \hat{\beta} )</th>
<th>( \tilde{\beta} )</th>
<th>( \beta^*(R_{max}, 1) )</th>
<th>( \delta^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female LFP</td>
<td>-13.83</td>
<td>-11.92</td>
<td>-8.35</td>
<td>1.44</td>
</tr>
<tr>
<td>Gender difference in LFP</td>
<td>-11.61</td>
<td>-14.66</td>
<td>-17.87</td>
<td>3.24</td>
</tr>
<tr>
<td>Female primary completion</td>
<td>14.79</td>
<td>-6.71</td>
<td>-19.40</td>
<td>( \delta &lt; 0 )</td>
</tr>
<tr>
<td>Gender difference in primary</td>
<td>1.21</td>
<td>-3.72</td>
<td>-6.27</td>
<td>( \delta &lt; 0 )</td>
</tr>
<tr>
<td>Female secondary completion</td>
<td>14.52</td>
<td>0.43</td>
<td>-9.69</td>
<td>0.05</td>
</tr>
<tr>
<td>Gender difference in secondary</td>
<td>0.48</td>
<td>-0.86</td>
<td>-1.77</td>
<td>( \delta &lt; 0 )</td>
</tr>
<tr>
<td>Gender attitude index</td>
<td>-0.03</td>
<td>-0.12</td>
<td>-0.20</td>
<td>( \delta &lt; 0 )</td>
</tr>
<tr>
<td>Gender attitudes: women</td>
<td>-0.02</td>
<td>-0.10</td>
<td>-0.18</td>
<td>( \delta &lt; 0 )</td>
</tr>
<tr>
<td>Gender attitudes: men</td>
<td>-0.04</td>
<td>-0.14</td>
<td>-0.23</td>
<td>( \delta &lt; 0 )</td>
</tr>
</tbody>
</table>

Where:

- \( \beta^* \) = implied causal impact of \( X \) on \( Y \) if \( \delta = 1 \)
- \( \delta^* \) = implied proportional selection coefficient under null
Within-Country Analysis
Within-Country Analysis: Afrobarometer Data

Gender languages account for between 10 and 90 percent of population of Chad, Kenya, Mauritania, Niger, Nigeria, S. Sudan, Uganda

Jakiela and Ozier (2018) Gendered Language, Slide 69
We pool Afrobarometer data from Kenya, Niger, Nigeria, Uganda:

<table>
<thead>
<tr>
<th>Survey Round</th>
<th>Kenya</th>
<th>Niger</th>
<th>Nigeria</th>
<th>Uganda</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round 3: 2005</td>
<td>1,261</td>
<td>0</td>
<td>2,120</td>
<td>2,345</td>
<td>5,726</td>
</tr>
<tr>
<td>Round 4: 2008</td>
<td>1,092</td>
<td>0</td>
<td>2,291</td>
<td>2,420</td>
<td>5,803</td>
</tr>
<tr>
<td>Round 5: 2011–2013</td>
<td>2,373</td>
<td>1,192</td>
<td>2,366</td>
<td>2,379</td>
<td>8,310</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>7,079</td>
<td>1,192</td>
<td>8,893</td>
<td>9,382</td>
<td>26,546</td>
</tr>
</tbody>
</table>

Identify grammatical gender structure for 99 percent of respondents

- Respondents speak 167 different African languages
Within-Country Analysis: IHDS Data

62 percent of the Indian population speaks a gender native language

- India Human Development Survey (IHDS) includes data on 75,966 household heads and spouses who speak 57 different languages

Jakiela and Ozier (2018) Gendered Language, Slide 71
Within-Country Analysis: Empirical Specifications

When we restrict the sample to women:

\[ Y_i = \alpha + \beta Gender_i + \zeta Z_i + \epsilon_i \]

where:

- **Gender_i** is an indicator for having a gender native language
- **X_i** is a vector of individual-level controls
  - Age, age\(^2\), religion indicators
- Regressions of Afrobarometer data also include country×round FEs
- \( \epsilon_i \) is a mean-zero error term
Within-Country Analysis: Empirical Specifications

When we include data on both women and men:

\[ Y_i = \alpha + \beta Gender_i + \eta Female_i + \theta Gender \times Female_i + \gamma_{country \times round} + \zeta Z_i + \epsilon_i \]

where:

- \( Gender_i \) is an indicator for having a gender native language
- \( Female_i \) is an indicator for being female
- \( Gender \times Female_i \) is a \( Gender_i \times Female_i \) interaction
- \( X_i \) is a vector of individual-level controls (age, religion, interactions)
- Regressions of Afrobarometer data also include country \( \times \) round FEs
- \( \epsilon_i \) is a mean-zero error term
Within-Country Analysis: Sets of eight coefficients

Jakiela and Ozier (2018)  Gendered Language, Slide 74
Within-Country Analysis: LFP

Coefficients on grammatical gender:
- Female Labor Force Participation
  - Africa without controls
  - Africa with controls
  - India without controls
  - India with controls

Gender Difference in LFP
- Africa without controls
- Africa with controls
- India without controls
- India with controls

Jakiela and Ozier (2018) Gendered Language, Slide 75
Within-Country Analysis: Primary Schooling

Jakiela and Ozier (2018) Gendered Language, Slide 76
Within-Country Analysis: Secondary Schooling

Jakiela and Ozier (2018) Gendered Language, Slide 77
Within-Country Analysis: Results

Labor Force Participation

Primary Completion

Secondary Completion

Jakiela and Ozier (2018) Gendered Language, Slide 78
Within-Country: Coefficient Stability

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>$\tilde{\beta}$</th>
<th>$\beta^*(R_{max}, 1)$</th>
<th>$\delta^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Afrobarometer Data from Kenya, Niger, Nigeria, and Uganda</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In labor force (women only)</td>
<td>-0.24</td>
<td>-0.18</td>
<td>-0.13</td>
<td>2.11</td>
</tr>
<tr>
<td>Female $\times$ in labor force (pooled)</td>
<td>-0.17</td>
<td>-0.11</td>
<td>-0.06</td>
<td>1.86</td>
</tr>
<tr>
<td>Completed primary (pooled)</td>
<td>-0.31</td>
<td>-0.22</td>
<td>-0.15</td>
<td>2.18</td>
</tr>
<tr>
<td>Female $\times$ primary (Table A8)</td>
<td>-0.12</td>
<td>-0.11</td>
<td>-0.10</td>
<td>4.64</td>
</tr>
<tr>
<td>Completed secondary (pooled)</td>
<td>-0.19</td>
<td>-0.16</td>
<td>-0.14</td>
<td>3.47</td>
</tr>
<tr>
<td>Female $\times$ secondary (Table A8)</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.06</td>
<td>6.01</td>
</tr>
<tr>
<td><strong>Panel B. India Human Development Survey III (IHDS) Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In labor force (women only)</td>
<td>-0.08</td>
<td>-0.07</td>
<td>-0.07</td>
<td>11.70</td>
</tr>
<tr>
<td>Female $\times$ in labor force (pooled)</td>
<td>-0.10</td>
<td>-0.08</td>
<td>-0.04</td>
<td>1.90</td>
</tr>
<tr>
<td>Completed primary (women only)</td>
<td>-0.14</td>
<td>-0.13</td>
<td>-0.12</td>
<td>12.14</td>
</tr>
<tr>
<td>Female $\times$ primary (pooled)</td>
<td>-0.13</td>
<td>-0.12</td>
<td>-0.11</td>
<td>13.19</td>
</tr>
<tr>
<td>Completed secondary (women only)</td>
<td>-0.03</td>
<td>-0.02</td>
<td>-0.02</td>
<td>7.20</td>
</tr>
<tr>
<td>Female $\times$ secondary (pooled)</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>25.89</td>
</tr>
</tbody>
</table>

Policy ramifications
Policy Implications: More Than Words

Gender matters when it shouldn’t.

• Bohren, Imas, and Rosenberg (2018a,b) show experimentally that randomizing the gender of the account name:
  ▶ Elicits differently-toned responses (using more opinion words) when the account posing the question is female-named;
  ▶ and elicits a lower subjective rating — fewer “upvotes” — when new users posting questions a female-named (though the pattern reverses with more established accounts).

• Boring, Ottoboni, and Stark (2016) show that students give higher evaluation ratings to instructors whom they perceive to be male.
  ▶ True even when (a) the instructors are actually female (online randomization by MacNell, Driscoll and Hunt 2014)
  ▶ and when (b) the male instructors produce worse learning outcomes (natural experiment analyzed by Boring 2017).

Jakiela and Ozier (2018) Gendered Language, Slide 81
Gender matters when it shouldn’t.

- Bohren, Imas, and Rosenberg (2018a,b) show experimentally that randomizing the gender of the account name:
  - Elicits differently-toned responses (using more opinion words) when the account posing the question is female-named;
  - and elicits a lower subjective rating — fewer “upvotes” — when new users posting questions a female-named (though the pattern reverses with more established accounts).
- Boring, Ottoboni, and Stark (2016) show that students give higher evaluation ratings to instructors whom they perceive to be male.
  - True even when (a) the instructors are actually female (online randomization by MacNell, Driscoll and Hunt 2014)
  - and when (b) the male instructors produce worse learning outcomes (natural experiment analyzed by Boring 2017).
Gender matters when it shouldn’t.

- Bohren, Imas, and Rosenberg (2018a,b) show experimentally that randomizing the gender of the account name:
  - Elicits differently-toned responses (using more opinion words) when the account posing the question is female-named;

- Boring, Ottoboni, and Stark (2016) show that students give higher evaluation ratings to instructors whom they perceive to be male.
  - True even when (a) the instructors are actually female (online randomization by MacNell, Driscoll and Hunt 2014)
  - and when (b) the male instructors produce worse learning outcomes (natural experiment analyzed by Boring 2017).
Policy Implications: More Than Words

Gender matters when it shouldn’t.

- Bohren, Imas, and Rosenberg (2018a,b) show experimentally that randomizing the gender of the account name:
  - Elicits differently-toned responses (using more opinion words) when the account posing the question is female-named;
  - and elicits a lower subjective rating — fewer “upvotes” — when new users posting questions a female-named (though the pattern reverses with more established accounts).

- Boring, Ottoboni, and Stark (2016) show that students give higher evaluation ratings to instructors whom they perceive to be male.
  - True even when (a) the instructors are actually female (online randomization by MacNell, Driscoll and Hunt 2014)
  - and when (b) the male instructors produce worse learning outcomes (natural experiment analyzed by Boring 2017).
Policy Implications: More Than Words

Gender matters when it shouldn’t.

• Bohren, Imas, and Rosenberg (2018a,b) show experimentally that randomizing the gender of the account name:
  ▶ Elicits differently-toned responses (using more opinion words) when the account posing the question is female-named;
  ▶ and elicits a lower subjective rating — fewer “upvotes” — when new users posting questions a female-named (though the pattern reverses with more established accounts).

• Boring, Ottoboni, and Stark (2016) show that students give higher evaluation ratings to instructors whom they perceive to be male.
Policy Implications: More Than Words

Gender matters when it shouldn’t.

- Bohren, Imas, and Rosenberg (2018a,b) show experimentally that randomizing the gender of the account name:
  - Elicits differently-toned responses (using more opinion words) when the account posing the question is female-named;
  - and elicits a lower subjective rating — fewer “upvotes” — when new users posting questions a female-named (though the pattern reverses with more established accounts).

- Boring, Ottoboni, and Stark (2016) show that students give higher evaluation ratings to instructors whom they perceive to be male.
  - True even when (a) the instructors are actually female (online randomization by MacNell, Driscoll and Hunt 2014)

Jakiela and Ozier (2018)  Gendered Language, Slide 81
Policy Implications: More Than Words

Gender matters when it shouldn’t.

- Bohren, Imas, and Rosenberg (2018a,b) show experimentally that randomizing the gender of the account name:
  - Elicits differently-toned responses (using more opinion words) when the account posing the question is female-named;
  - and elicits a lower subjective rating — fewer “upvotes” — when new users posting questions a female-named (though the pattern reverses with more established accounts).

- Boring, Ottoboni, and Stark (2016) show that students give higher evaluation ratings to instructors whom they perceive to be male.
  - True even when (a) the instructors are actually female (online randomization by MacNell, Driscoll and Hunt 2014)
  - and when (b) the male instructors produce worse learning outcomes (natural experiment analyzed by Boring 2017).
Policy Implications: More Than Words

Interventions can leverage the salience of gender.
Policy Implications: More Than Words

Interventions can leverage the salience of gender.

- Porter and Serra (2018) show that having a female role model visit a Principles of Economics class makes female students more likely to take Intermediate Micro, and to consider majoring in economics (no impact on male students).
Policy Implications: More Than Words

Phiona Mutesa

Policy Implications: More Than Words

Interventions can leverage the salience of gender.

- Porter and Serra (2018) show that having a female role model visit a Principles of Economics class makes female students more likely to take Intermediate Micro, and to consider majoring in economics (no impact on male students).

- Riley (2018) shows that watching *Queen of Katwe* causes Ugandan secondary school students to perform better on a mathematics examination, with largest effects for female and lower-ability students.
Interventions can leverage the salience of gender.

- Porter and Serra (2018) show that having a female role model visit a Principles of Economics class makes female students more likely to take Intermediate Micro, and to consider majoring in economics (no impact on male students).

- Riley (2018) shows that watching *Queen of Katwe* causes Ugandan secondary school students to perform better on a mathematics examination, with largest effects for female and lower-ability students.

- sadietannerconference.org
  “You can’t be what you can’t see” - Dr. Joycelyn Elders
Policy Implications: More Than Words

Interventions can address misperceptions about common beliefs.

- Bursztyn, González, and Yanagizawa-Drott (2018) show that
Interventions can address misperceptions about common beliefs.

- Bursztyn, González, and Yanagizawa-Drott (2018) show that
  - Men in Saudi Arabia *believe* that other men are less supportive of female labor force participation than they really are;
Policy Implications: More Than Words

Interventions can address misperceptions about common beliefs.

- Bursztyn, González, and Yanagizawa-Drott (2018) show that
  - Men in Saudi Arabia believe that other men are less supportive of female labor force participation than they really are;
  - Correcting beliefs experimentally makes men more willing to forgo income so that wives can participate in online job-matching;

- Patnaik (2018) shows that the Quebec Parental Insurance Program’s “daddy-only” label for some parts of parental leave
  - increased fathers’ use of parental leave (53 percentage pts, 200 pct);
  - also increased fathers’ long-term share of household and child-rearing responsibilities, increasing mothers’ labor supply as well.
Interventions can address misperceptions about common beliefs.

• Bursztyn, González, and Yanagizawa-Drott (2018) show that
  ▶ Men in Saudi Arabia believe that other men are less supportive of female labor force participation than they really are;
  ▶ Correcting beliefs experimentally makes men more willing to forgo income so that wives can participate in online job-matching;
  ▶ Months later, this increases the likelihood that women have participated in a job interview outside the home.
Interventions can address misperceptions about common beliefs.

- Bursztyn, González, and Yanagizawa-Drott (2018) show that
  - Men in Saudi Arabia believe that other men are less supportive of female labor force participation than they really are;
  - Correcting beliefs experimentally makes men more willing to forgo income so that wives can participate in online job-matching;
  - Months later, this increases the likelihood that women have participated in a job interview outside the home.

- Patnaik (2018) shows that that the Quebec Parental Insurance Program’s “daddy-only” label for some parts of parental leave
Interventions can address misperceptions about common beliefs.

- Bursztyn, González, and Yanagizawa-Drott (2018) show that
  - Men in Saudi Arabia believe that other men are less supportive of female labor force participation than they really are;
  - Correcting beliefs experimentally makes men more willing to forgo income so that wives can participate in online job-matching;
  - Months later, this increases the likelihood that women have participated in a job interview outside the home.

- Patnaik (2018) shows that the Quebec Parental Insurance Program’s “daddy-only” label for some parts of parental leave
  - increased fathers’ use of parental leave (53 percentage pts, 200 pct);
Policy Implications: More Than Words

Interventions can address misperceptions about common beliefs.

- Bursztyn, González, and Yanagizawa-Drott (2018) show that
  ▶ Men in Saudi Arabia believe that other men are less supportive of female labor force participation than they really are;
  ▶ Correcting beliefs experimentally makes men more willing to forgo income so that wives can participate in online job-matching;
  ▶ Months later, this increases the likelihood that women have participated in a job interview outside the home.

- Patnaik (2018) shows that that the Quebec Parental Insurance Program’s “daddy-only” label for some parts of parental leave
  ▶ increased fathers’ use of parental leave (53 percentage pts, 200 pct);
  ▶ also increased fathers’ long-term share of household and child-rearing responsibilities, increasing mothers’ labor supply as well.
Portrait of the Acting CE as a Young[er] Man

Source: @Shanta_WB
Conclusions

We characterize the grammatical gender structure of most of the world’s living languages, accounting for 99 percent of the population.

We present cross-country evidence that gender languages:

- Predict lower female labor force participation, gender attitudes

We present within-country evidence that gender languages:

- Predict lower female labor force participation, schooling

Languages have inherent cultural value, but they change over time; some changes result from policy choices (e.g. Académie Francaise)

- Our results suggest that linguistic choices - and many other nudges - should be seen as policy
Thank you!
Additional Slides
### Marginal Impact of Dichotomous Gender Categories

<table>
<thead>
<tr>
<th>Specification</th>
<th>LFP&lt;sub&gt;f&lt;/sub&gt;</th>
<th>LFP&lt;sub&gt;f&lt;/sub&gt; - LFP&lt;sub&gt;m&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion (any) gender</td>
<td>-6.66±2.54</td>
<td>4.29±1.65</td>
</tr>
<tr>
<td>Proportion dichotomous gender</td>
<td>-10.58±4.78</td>
<td>-23.44±3.54</td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Geography Controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>178</td>
<td>178</td>
</tr>
<tr>
<td>R²</td>
<td>0.19</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Robust standard errors clustered by most widely spoken language in all specifications. LFP<sub>f</sub> is the percentage of women in the labor force, measured in 2011. LFP<sub>m</sub> - LFP<sub>f</sub> is the difference between male and female labor force participation in 2011. Strong gender languages are those that partition the space of nouns into two gender categories, masculine and feminine. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina et al. (2013) measure of suitability for the plough.

→ Return to robustness checks

### Robustness to Potentially Endogenous Controls

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>LFP&lt;sub&gt;f&lt;/sub&gt;</th>
<th>LFP&lt;sub&gt;f&lt;/sub&gt; - LFP&lt;sub&gt;m&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specification:</strong></td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Proportion speaking gender language</td>
<td>-6.66</td>
<td>-10.42</td>
</tr>
<tr>
<td></td>
<td>(2.80)</td>
<td>(2.84)</td>
</tr>
<tr>
<td></td>
<td>[p &lt; 0.001]</td>
<td>[p &lt; 0.001]</td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country-Level Geography Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>176</td>
<td>176</td>
</tr>
<tr>
<td>R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.57</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Robust standard errors clustered by most widely spoken language in all specifications. LFP<sub>f</sub> is the percentage of women in the labor force, measured in 2011. LFP<sub>m</sub> - LFP<sub>f</sub> is the difference between male and female labor force participation in 2011. Strong gender languages are those that partition the space of nouns into two gender categories, masculine and feminine. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina et al. (2013) measure of suitability for the plough. Bad controls are log GDP per capita (in 2011), log population (in 2011), and the percentage Catholic, Protestant, other Christian, Muslim, and Hindu (taken from Alesina et al. 2013), and an indicator for former communist countries.

→ Return to robustness checks
### Omitting Major World Languages

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>LFP&lt;sub&gt;f&lt;/sub&gt;</th>
<th>LFP&lt;sub&gt;f&lt;/sub&gt; – LFP&lt;sub&gt;m&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Omitted Language:</strong></td>
<td>Arabic OLS (1)</td>
<td>Arabic OLS (4)</td>
</tr>
<tr>
<td></td>
<td>English OLS (2)</td>
<td>English OLS (5)</td>
</tr>
<tr>
<td></td>
<td>Spanish OLS (3)</td>
<td>Spanish OLS (6)</td>
</tr>
<tr>
<td>Proportion speaking gender language</td>
<td>-6.18 (3.56)</td>
<td>-9.09 (3.52)</td>
</tr>
<tr>
<td></td>
<td>-12.33 (3.84)</td>
<td>-15.31 (3.59)</td>
</tr>
<tr>
<td></td>
<td>-10.10 (3.87)</td>
<td>-11.31 (3.39)</td>
</tr>
<tr>
<td></td>
<td>[0.085]</td>
<td>[0.011]</td>
</tr>
<tr>
<td></td>
<td>[0.002]</td>
<td>[p &lt; 0.001]</td>
</tr>
<tr>
<td></td>
<td>[0.010]</td>
<td>[0.001]</td>
</tr>
<tr>
<td>Continent Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country-Level Geography Controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>159</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>167</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>160</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.21</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>0.34</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Robust standard errors are clustered by the most widely spoken language in all specifications; they are reported in parentheses. P-values are reported in square brackets. \( \text{LFP}_f \) is the percentage of women in the labor force, measured in 2011. \( \text{LFP}_f - \text{LFP}_m \) is the difference between male and female labor force participation in 2011. Geography controls are the percentage of land area in the tropics or subtropics, average yearly precipitation, average temperature, an indicator for being landlocked, and the Alesina et al. (2013) measure of suitability for the plough.

→ Return to robustness checks
Secondary School Completion by Continent

Return to secondary education figure
Comparing country-level WALS data to full data

Two measures of the fraction of a country speaking a gender language as their native language

→ Return to distribution of gender languages
Cross-Country Analysis: Measurement Error

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Naïve OLS CI</th>
<th>Imbens-Manski CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female labor force participation</td>
<td>$[-18.533, -5.305]$</td>
<td>$[-18.467, -5.013]$</td>
</tr>
<tr>
<td>Female primary school completion</td>
<td>$[-15.431, 2.010]$</td>
<td>$[-16.221, 1.673]$</td>
</tr>
<tr>
<td>Gender difference in primary school completion</td>
<td>$[-8.003, 0.559]$</td>
<td>$[-8.446, 0.432]$</td>
</tr>
<tr>
<td>Female secondary school completion</td>
<td>$[-6.901, 7.769]$</td>
<td>$[-8.261, 7.327]$</td>
</tr>
<tr>
<td>Gender difference in secondary school completion</td>
<td>$[-5.510, 3.799]$</td>
<td>$[-5.401, 3.746]$</td>
</tr>
<tr>
<td>Gender attitudes index</td>
<td>$[-0.193, -0.045]$</td>
<td>$[-0.194, -0.047]$</td>
</tr>
<tr>
<td>Gender attitudes index among women</td>
<td>$[-0.173, -0.022]$</td>
<td>$[-0.173, -0.023]$</td>
</tr>
<tr>
<td>Gender attitudes index among men</td>
<td>$[-0.214, -0.063]$</td>
<td>$[-0.215, -0.064]$</td>
</tr>
</tbody>
</table>

For each outcome, the naïve confidence interval comes from the associated regression in a previous table. The Imbens-Manski worst-case confidence interval is calculated by finding the minimum and maximum possible point estimates of the relevant coefficient based on the interval nature of the dataset (without complete data on the grammatical structure of all languages, the right-hand-side variable—the fraction of a country’s population speaking a gender language—is only observed up to an interval in some cases), then by tightening the confidence interval for correct coverage following Imbens and Manski (2004).