


Decomposing culture: an analysis of gender, language, and labor supply in the household

Victor Gay¹ · Daniel L. Hicks² · Estefania Santacreu-Vasut ³ · Amir Shoham⁴

Received: 24 August 2016 / Accepted: 27 March 2017
© Springer Science+Business Media New York 2017

Abstract Despite broad progress in closing many dimensions of the gender gap around the globe, recent research has shown that traditional gender roles can still exert a large influence on female labor force participation, even in developed economies. This paper empirically analyzes the role of culture in determining the labor market engagement of women within the context of collective models of household decision making. In particular, we use the epidemiological approach to study the relationship between gender in language and labor market participation among married female immigrants to the U.S. We show that the presence of gender in language can act as a marker for culturally acquired gender roles and that these roles are important determinants of household labor allocations. Female immigrants who speak a language with sex-based grammatical rules exhibit lower labor force participation, hours worked, and weeks worked. Our strategy of isolating one component of culture reveals that roughly two thirds of this relationship can be explained by correlated cultural factors, including the role of bargaining power in the household, and the impact of ethnic enclaves and that at most one third is potentially explained by language having a causal impact.

Electronic supplementary material The online version of this article (doi:[10.1007/s11150-017-9369-x](https://doi.org/10.1007/s11150-017-9369-x)) contains supplementary material, which is available to authorized users.

✉ Estefania Santacreu-Vasut
santacreuvasut@essec.edu

¹ Department of Economics, The University of Chicago, 5757S. University Ave, Chicago, IL 60637, USA

² Department of Economics, University of Oklahoma, 308 Cate Center Drive, Norman, OK 73019, USA

³ Department of Economics, ESSEC Business School and THEMA, 3 Avenue Bernard Hirsch, 95021 Cergy-Pontoise, France

⁴ Temple University and COMAS, 1801N. Broad St, Philadelphia, PA 19122, USA

Keywords Language · Gender gap · Bargaining power in the household · Labor force participation · Immigrants

JEL Classification F22 · J16 · Z13

1 Introduction

The labor force participation rate of married women has increased spectacularly during the second half of the 20th century, both in the U.S. and around the world (Goldin 2014). Nevertheless, wide differences across countries remain, with the gender gap closing faster in some countries than others, and a number of developing nations, such as India, lagging behind (Field et al. 2016). Among the factors used to explain these differences, the role of cultural norms has been shown to play a crucial role (Fernández 2007; Fernández and Fogli 2009; Alesina et al. 2013; Farré and Vella 2013; Fernández 2013). In particular, scholars have begun employing the epidemiological approach to distinguish the impact of cultural factors from that of other institutional forces (Fernández and Fogli 2009; Fernández 2011).¹

When analyzing married couples, both individual and spousal culture are thought to play an important role in explaining labor market decisions. Early collective models of labor supply emphasize the importance of taking into account the distribution of bargaining power within the household (Chiappori et al. 2002, Blundell et al. 2007). Recently however, these models have been extended to investigate whether their main predictions depend not only on bargaining power, but also on the surrounding cultural context (Oreffice 2014) as well as on social and institutional constraints related to traditional gender roles (Field et al. 2016). These norms may influence women to allocate more labor to household production relative to the formal market at either the intensive or extensive margin.

This paper studies the impact of gender roles on the labor force participation of married female immigrants in the U.S. within a collective labor supply framework. We explore culturally acquired gender roles as those embodied by the presence of grammatical gender in language spoken. Recent work has documented correlations between linguistic structure and individual behavior (Lupyan and Dale 2010; Chen 2013; Ladd et al. 2015) with the presence and intensity of gender in grammatical structures correlating with gender gaps in compensation and promotions, division of household labor, educational attainment, and political empowerment (Givati and Troiano 2012; Santacreu-Vasut et al. 2013, 2014; Hicks et al. 2015; Mavisakalyan 2015; van der Velde et al. 2015; Davis and Reynolds 2016).

The mechanisms underlying these associations are largely unresolved. Perhaps the most compelling question, whether language may causally influence behavior, remains a subject of debate.² Linguistic correlations could reflect historically

¹ Gay et al. (2016) provides an in-depth discussion of the epidemiological approach in this context.

² Theoretically language structures could influence preference formation, information processing, categorization of social reality, and the salience of certain categories of words. These impacts could alter a speaker's decision making and behavior (Mavisakalyan and Weber 2016).

acquired cultural norms of behavior that became codified in language or language itself could represent an institutional force, influencing and perpetuating a set of behaviors.

From a methodological point of view, because language is observed at the individual level, it is possible to isolate its effect from the influence of aggregate correlated historical, cultural, and biological forces. This is an exercise we undertake in this analysis following the epidemiological approach. Furthermore, because language acts a cultural marker for gender roles, we show that this barometer of norms allows us to examine labor market predictions of the collective household model in the presence or absence of these cultural influences. This yields new insights regarding their relative influence.

To do this, we examine labor market outcomes for nearly half a million married female immigrants to the U.S. aged 25 to 49 in the American Community Surveys (ACS) between 2005 and 2015. In our analysis, we focus on married, foreign born individuals, who speak a language other than English in the home. With the assistance of linguists and drawing on information contained in the World Atlas of Linguistic Structures database (WALS), we assign to these speakers consistent measures of the presence and intensity of gender in the grammatical structure of their language.

There are several novel empirical advances to the approach taken in this paper. First, this analysis relies on a detailed set of microeconomic data which provides a rich set of covariates and substantial variation to control for many potential confounding factors. These include many of the individual, spousal, and household level variables analyzed in the past, as well as measures which may influence the decision to migrate in the country of origin and measures of the location to which the immigrant has moved. A second advantage is reliance on the epidemiological approach in this setting which helps ameliorate a fundamental problem of identification. By using a restrictive fixed effects strategy, including country fixed effects, we are able to capture a wide array of unobservable cultural forces and obtain identification using variation in language spoken from individuals of the same country of origin or ancestry.

The takeaways from this exercise are several. First, we show that married female immigrants speaking a language with sex-based distinctions in its grammar are less likely to participate in the labor market. This is true even after controlling for observable characteristics such as traditional household measures, husband characteristics, and bargaining power measures, as well as when controlling for a vast set of unobservable cultural forces through country of origin fixed effects.³ When we empirically decompose the relationship between gender in language and gendered behavior in this manner, it suggests that roughly two thirds of this relationship can be explained by correlated cultural factors, with about one third potentially explained by language having a causal impact.

Second, using language as a measure of culturally acquired gender roles allows us to speak to and test the role of bargaining power within the household relative to the

³ One of the challenges of the literature on culture and economic behavior has been to measure culture. An advantage of using language over existing alternatives is that language is defined at the individual level and is not an outcome measure. It allows us to control for country of origin characteristics including time varying ones.

impact of cultural norms. We demonstrate that both are important, and that their impacts are distinct from one another.

Third, focusing on language spoken allows to also study the behavior of female immigrants in both linguistically homogeneous and heterogeneous couples. We exploit linguistic heterogeneity within the household to show that the presence of gender in language has an association with a wife's behavior both when husbands and wives speak a gendered language, and when they speak languages with different structures. Our findings suggest that while there is an independent effect on a wife's behavior when she alone has a gender marked language, gender marking in the husbands language enhances this effect.

Fourth, using language as a measure of gender roles but recognizing it as a network technology allows us to examine the role that ethnic enclaves play in influencing female labor force participation. In theory, enclaves may improve labor market outcomes by providing information about formal jobs and reducing social stigma on employment. At the same time, enclaves are likely to reinforce immigrant language usage and thus may enhance the impact of gender in language or provide isolation from U.S. norms. We present evidence implying that the latter effect is present, which suggests that the impact of language is stronger when it is shared with the surrounding community.⁴

Finally, and as part of our robustness checks, we also show that some forms of gender roles are not present for unmarried women suggesting that they may be dormant among unmarried women where the pressure to not work, to raise children, and to provide household goods may be weaker.

These findings speak to several literatures. Mavisakalyan and Weber (2016) review of the nascent field of linguistic relativity and economics and point out that the mechanisms behind the associations between language and economic behavior remain largely unresolved. Despite the current study's ability to account for both time-invariant and time-variant country of origin factors in a more rigorous manner than previous analyses, the correlation between language and behavior remains significant and negative. This means that while we can demonstrate that most of this association can be explained by language as a cultural market for correlated gender norms, we cannot formally disprove that the behavioral channel may explain some of the remaining correlation (Roberts and Winters 2013; Hicks et al. 2015; Roberts et al. 2015).⁵

Our research also directly contributes to literature that investigates whether the impact of intra-household bargaining power on the labor supply depends on culture. In particular, our results imply that the standard prediction that the spouse with higher bargaining power will substitute labor for leisure due to an income effect as in Orefice (2014) applies only to native born couples in the U.S., and to immigrant couples with gender roles similar to the U.S. Interestingly, our results confirm that in couples coming from countries classified as exhibiting strong traditional gender roles, the influence of bargaining power on household collective labor supply may be

⁴ This result is not solely attributable to selection since it holds for women married prior to migration—sometimes referred to as “tied women”.

⁵ See Everett (2013) for a review of the linguistics literature.

culturally dependent and that some of these standard predictions may not hold for these groups.

While Orefice (2014) focuses on the intensive labor supply among couples where both spouses are working, we focus mainly on the extensive margin and how bargaining power influences female labor force participation. Field et al. (2016) study the extensive margin of the labor supply among married female in India and model the impact of bargaining power distribution within the household with social constraints related to traditional gender norms. Our findings complement these studies by empirically showing that the married female immigrants with stronger bargaining power are more likely to work, not less, and that the impact of bargaining power is reinforced when traditional gender roles are embodied in the language spoken.

This paper is structured as follows. Section 2 presents summary statistics for both the ACS and linguistic data and details the empirical strategy. Section 3 presents the empirical results—decomposing the impact of language from other cultural influences, using gender in language as a cultural marker to study labor supply within the household and presenting a wide range of robustness checks. Section 4 concludes.

2 Methodology

2.1 Data

2.1.1 Economic and demographic data

We combine data from several sources in our analysis. First, we obtain detailed demographic and economic data for the immigrant population to the U.S. from the American Community Surveys (ACS) 1% samples from 2005 to 2015 (Ruggles et al. 2015).

We restrict the analysis to female respondents aged 25 to 49, born abroad from non-American parents, living in married-couple family households in which the husband is present, and who report speaking a language other than English in the home. Moreover, we only keep respondents that report uniquely identifiable countries of birth and languages, and for which we have information on the grammatical structure of the language reported. Online Appendix A.1.1 provides an exhaustive list of these sample restrictions, and details the precise process taken to construct the regression sample. It also provides precise definitions for all variables employed.⁶

The main dependent variable used throughout the analysis is a measure of labor market engagement, labor force participation, which is defined as an indicator equal to one if the respondent is either employed or actively looking for a job.⁷ As robustness check, we also analyze other labor market outcomes such as yearly

⁶ Formal checks suggest that the regression sample is not biased by the availability of the country of birth and the linguistic data. Online Appendix Table C.2 provides summary statistics comparable to those in Table 1 without dropping the respondents for which the country of birth or language are not precisely identifiable, which is about 6.78% of the uncorrected regression sample. The last column of this table demonstrates that data constraints—i.e., needing to know language spoken and country of birth—do not meaningfully bias the sample in any way along these observables.

⁷ This is the **LABFORCE** variable in the ACS (Ruggles et al. 2015). See Online Appendix A.2.1 for more details.

weeks worked and usual weekly hours worked, both including and excluding zeros, to capture both the extensive and intensive margins of the labor supply.⁸

Table 1 presents summary statistics for the demographic and economic characteristics of the respondents in the regression sample. On average, the typical married women in our sample is in her upper 30s, immigrated around 15 years before, and has over 12 years of education. 60% of these respondents participate in the labor force, with 55% reporting formal employment. As can be seen from the last column of Table 1, there is a large difference in means between the labor market outcomes for sex-based speakers and non sex-based speakers, with the former group exhibiting far lower levels of economic engagement. There is also sizeable variation in English proficiency, as well as in racial and ethnic composition for this sample. Mean household income is just over \$60,000, and the average duration of the current marriage is 12 years. Many households have children, with the mean being almost two. Overall, the population studied contains rich variation, with over 480,000 adult female immigrants originating from 135 countries and speaking 63 different languages. Online Appendix Table C.1 provides the distribution of languages spoken in the regression sample and highlights the extensive variation in languages spoken by immigrants to the U.S. Moreover, while they are not the primary groups of interest in this analysis, we further provide similar summary statistics for the respondents' husbands in Online Appendix Table C.3, as well as various within-household gender gap measures in Online Appendix Table C.4, both of which are used in subsequent analysis.

2.1.2 Linguistic data

Next, we follow Gay et al. (2013) and Hicks et al. (2015), and assign to each language measures that quantify the presence and frequency of gender distinctions in its grammatical rules. We construct these measures using information compiled by linguists in the World Atlas of Language Structures (WALS, Dryer and Haspelmath 2011). We expand the original WALS dataset in collaboration with linguists for several additional languages, making the sample more representative.⁹

Our primary measure of gender in language is an indicator for whether a language employs a grammatical gender system based on biological sex (**SB**). We also investigate gender distinctions in other features of the grammatical structure. For instance, languages with only a male and a female gender force speakers to make more sex-based distinctions than those which include a neuter gender. **NG** is an indicator variable equal to one for languages with exactly two genders, and equal to zero otherwise. Similarly, there is heterogeneity across languages in the presence and quantity of gendered personal pronouns. This feature is given by the variable **GP**, which captures rules related to gender agreement with pronouns. Finally, some

⁸ The yearly weeks worked correspond to the **WKSWORK2** variable in the ACS (Ruggles et al. 2015). Because this variable is given in intervals, we define yearly weeks worked as the midpoint of those intervals. The usual weekly hours worked correspond to the **UHRSWORK** variable in the ACS (Ruggles et al. 2015). Online Appendix A.2.1 provides additional details.

⁹ The languages for which some variables were compiled are detailed in Online Appendix B. For robustness, we have also verified that the exclusion of our newly assigned languages in favor of the original WALS set of languages does not alter the main findings of the analysis.

Table 1 Summary statistics: Female immigrants, married, spouse present, aged 25–49

	A. Individual characteristics					
	Mean	S.d.	Min.	Max.	Obs.	SBI–SB0
Age	37.7	6.6	25	49	480,618	–1.0***
Years since immigration	14.8	9.3	0	50	480,618	–0.0
Age at immigration	22.8	8.9	0	49	480,618	–1.0***
Educational attainment						
Current student	0.06	0.24	0	1	480,618	–0.02***
Years of schooling	12.4	3.7	0	17	480,618	–1.3***
No schooling	0.05	0.21	0	1	480,618	0.00*
Elementary	0.12	0.32	0	1	480,618	0.10***
High school	0.37	0.48	0	1	480,618	0.10***
College	0.46	0.50	0	1	480,618	–0.19***
Race and ethnicity						
Asian	0.29	0.46	0	1	480,618	–0.65***
Black	0.02	0.14	0	1	480,618	0.01***
White	0.47	0.50	0	1	480,618	0.41***
Other	0.21	0.41	0	1	480,618	0.24***
Hispanic	0.53	0.50	0	1	480,618	0.63***
Ability to speak English						
Not at all	0.11	0.32	0	1	480,618	0.08***
Not well	0.24	0.43	0	1	480,618	0.02***
Well	0.25	0.43	0	1	480,618	–0.08***
Very well	0.40	0.49	0	1	480,618	–0.02***
Labor market outcomes						
Labor participant	0.60	0.49	0	1	480,618	–0.10***
Employed	0.55	0.50	0	1	480,618	–0.12***
Yearly weeks worked (excl. 0)	44.16	13.2	7	51	302,653	–1.6***
Yearly weeks worked (incl. 0)	27.18	23.9	0	51	480,618	–5.8***
Weekly hours worked (excl. 0)	36.57	11.4	1	99	302,653	–1.8***
Weekly hours worked (incl. 0)	22.51	19.9	0	99	480,618	–5.1***
Labor income (thds.)	25.3	28.5	0.0	536.5	282,057	–8.8***
B. Household characteristics						
Years since married	12.4	7.5	0	39	352,059	0.14***
Number of children aged <5	0.42	0.65	0	6	480,618	0.04***
Number of children	1.87	1.26	0	9	480,618	0.38***
Household size	4.26	1.62	2	20	480,618	0.39***
Household income (thds.)	61.0	59.6	–31.0	1530.3	480,618	–20.8***

Notes: The summary statistics are computed using sample weights (PERWT) provided in the ACS (Ruggles et al. 2015), except those for household characteristics, which are computed using household weights (HHWT). The last column reports the estimate β from regressions of the type $X_i = \alpha + \beta \text{SBI} + \varepsilon$, where X_i is an individual level characteristic. Robust standard errors are not reported. See Online Appendix A for details on variable sources and definitions

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level

languages assign gender due to semantic reasons only, while others assign gender due to both semantic and formal reasons, making gender more recurrent in the latter case. This feature is given by the variable **GA**, which captures the rules for gender assignment.¹⁰

Finally, we employ a measure of grammatical gender intensity similar to the one built by Gay et al. (2013). This measure captures how many of the above features are present in a language. It is defined as:

$$\text{Intensity} = \text{SB} \times (\text{GP} + \text{GA} + \text{NG}),$$

where **Intensity** is a categorical variable that ranges from 0 to 3.¹¹ The **Intensity** measure allows us to capture the ranking of intensity of female/male distinctions in the grammatical rules as follows. If a language has a sex-based gender system, its intensity score is the sum of its scores on the three other gender features. Hence, a strictly positive score captures languages that have strong female/male distinctions. Table 2 provides summary statistics for the language variables contained in the regression sample.

There is substantial linguistic heterogeneity in terms of grammatical gender in the sample. We demonstrate the robustness of our main findings across all of these measures, with additional checks presented in Online Appendix Table C.7.

2.2 Empirical strategy

Our empirical strategy follows the epidemiological approach to culture Fernández and Fogli 2009; Fernández 2011; Blau et al. 2011; Blau and Kahn 2015). This approach compares outcomes across immigrants with varying geographical origins, but living in a common institutional, legal, and social environment, thereby allowing to separate cultural influences acquired prior to migration from confounding institutional forces. In the baseline specifications, we include a set of controls that are common in analyzing decision making in the collective household framework. Moreover, to help isolate the role of language from other cultural forces, we include fixed effects by country of birth of the respondents. These fixed effects allow us to obtain identification off heterogeneity in the structure of languages spoken across immigrants from the same country of birth. We generate our core results by estimating the following specification:

$$Y_{ijklst} = \alpha + \beta \text{SB}_{il} + \gamma' \mathbf{X}_{ij} + \delta' \mathbf{W}_c + \eta' \mathbf{S}_s + \theta' \mathbf{Z}_t + \varepsilon_{ijklst}, \quad (1)$$

where Y_{ijklst} is a measure of labor market participation. Subscript i indexes respondents; l , languages spoken; j , households; c , countries of birth; s , states of residence; t , ACS survey years. \mathbf{X}_{ij} corresponds to the characteristics of respondent i in household j . This vector contains the following variables: age, age squared, 5-years age group indicators, race indicators, a Hispanic indicator, age at immigration to the U.S., years since in the U.S., a student indicator, years of schooling, level of English

¹⁰ More detailed definitions of these individual measures can be found in Hicks et al. (2015).

¹¹ This measure should not be taken as measuring absolute intensity but rather as a ranking of relative intensity across languages grammar. The discussion of the measure is extended in Online Appendix discussion of Table C.7.

Table 2 Summary statistics, language variables: Female immigrants, married, spouse present, aged 25–49

	Definition	Mean	S.d.	Min.	Max.	Obs.
SB	Sex-based	0.83	0.37	0	1	480,618
NG	Number of genders	0.72	0.45	0	1	480,618
GA	Gender assignment	0.76	0.43	0	1	480,618
GP	Gender pronouns	0.57	0.50	0	1	478,883
Intensity	$(GP + GA + NG) \times SB$	2.04	1.21	0	3	478,883

Notes: The summary statistics are computed using sample weights (PERWT) provided in the ACS (Ruggles et al. 2015)

proficiency, decade of immigration indicators, number of children aged less than 5 years old in the household, and household size. W_c corresponds to a vector of indicators for the respondent's country of birth c , S_s corresponds to a vector of indicators for the respondent's state of residence s , and Z_t corresponds to a vector of indicators for the ACS survey year t . Online Appendices A.2.2 and A.2.3 provide details on how these variables are constructed.

To help the interpretation of regression coefficients, we additionally report in all regression tables the average of the outcome variable for the relevant sample. Moreover, to facilitate the comparison of magnitudes across coefficients from a given regression, we also report the coefficients when standardizing continuous variables between zero and one—we keep indicator variables unstandardized for ease of interpretation.

3 Empirical results

3.1 Decomposing language from other cultural influences

Table 3 contains our baseline empirical analysis. In the first column, we examine the naïve association between gender in language and labor force participation. These results are naïve in the sense that we are not yet accounting for any confounding factors, and simply represent the difference in means across groups. The raw correlation implies that married immigrant women who speak a language that has a sex-based gender structure are 10 percentage points less likely to participate in the formal labor market. This is about 17% of the average labor force participation rate for the full regression sample.

Moving across columns in the table, the analysis includes an increasingly stringent set of controls for both observable and unobservable factors which may impact a woman's decision to participate in the labor market. The specification reported in column (2) controls for the individual characteristics of the vector X_{ij} described in Section 2.2. After this inclusion, the impact of the SB variable remains statistically significant and negative, while its magnitude decreases. This suggests that part of the correlation between language and labor force participation emanates from the influence of individual characteristics correlated with both language structure and behavior. The decline in the magnitude of the coefficient is largely driven by the

Table 3 Gender in language and economic participation, individual level; Female immigrants, married, spouse present, 25–49 (2005–2015)

Dependent variable:	Labor force participation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sex-based	-0.101*** [0.002]	-0.069*** [0.003]		-0.045*** [0.004]	-0.027*** [0.007]	-0.024*** [0.007]	-0.021*** [0.008]
β -coef.				-0.045			
COB LFP			0.002*** [0.000]	0.002*** [0.000]			
β -coef.			0.041	0.039			
Past migrants LFP			0.005*** [0.000]	0.005*** [0.000]			
β -coef.			0.063	0.062			
Respondent char.							
Age	No	Yes	Yes	Yes	Yes	Yes	Yes
Race and ethnicity	No	Yes	Yes	Yes	Yes	Yes	Yes
Education	No	Yes	Yes	Yes	Yes	Yes	Yes
Years since immigr.	No	Yes	Yes	Yes	Yes	Yes	Yes
Age at immigr.	No	Yes	Yes	Yes	Yes	Yes	Yes
Decade of immigr.	No	Yes	Yes	Yes	Yes	Yes	Yes
English prof.	No	Yes	Yes	Yes	Yes	Yes	Yes
Household char.							
Children aged <5	No	Yes	Yes	Yes	Yes	Yes	Yes
Household size	No	Yes	Yes	Yes	Yes	Yes	Yes
State	No	Yes	Yes	Yes	Yes	Yes	Yes

Table 3 continued

Dependent variable:	Labor force participation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
COB char.							
Geography	No	No	Yes	Yes	No	No	No
Fertility	No	No	Yes	Yes	No	No	No
Migration rate	No	No	Yes	Yes	No	No	No
Education	No	No	Yes	Yes	No	No	No
GDP per capita	No	No	Yes	Yes	No	No	No
Common language	No	No	Yes	Yes	No	No	No
Genetic distance	No	No	Yes	Yes	No	No	No
COB FE	No	No	No	No	Yes	Yes	Yes
COB × decade mig. FE	No	No	No	No	No	Yes	Yes
County FE	No	No	No	No	No	No	Yes
Observations	480,618	480,618	451,048	451,048	480,618	480,618	382,556
R ²	0.006	0.110	0.129	0.129	0.132	0.138	0.145
Mean	0.60	0.60	0.60	0.60	0.60	0.60	0.60
SB residual variance	0.150	0.098		0.049	0.013	0.012	0.012

Notes: The estimates are computed using sample weights (PERWT) provided in the ACS (Ruggles et al. 2015). □-coef. denotes the coefficients from using standardized values of continuous variables in the regressions

COB country of birth, LFP female labor participation, FE fixed effects. See Online Appendix A for details on variable sources and definitions

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level

inclusion of controls for race, not by the addition of the other respondent or household characteristics.¹² Note that while we only report the coefficient of interest, the estimates on the other variables have the expected sign and magnitude. For instance, more educated respondents are more likely to be in the labor force, and those with more children aged less than 5 years old are less likely to be in the labor force. The coefficients for these variables are shown in Online Appendix Table C.6.

Table 3 also reports a measure of residual variance in the SB variable in the last row. This measure gives a sense of the variation left in the independent variable that is used in the identification of the coefficients after its correlation with other regressors has been accounted for. For instance, while the initial variance in the SB variable is 0.150—see column (1)—, adding the rich set of household and respondent characteristics in column (2) removes roughly one third of its variation.

As the historical development of languages was intertwined with cultural and biological forces, the observed associations in column (2) could reflect the impact of language, the influence of environmental gender norms acquired prior to migration through other channels, or both. As a first step in disentangling these potential channels, column (3) controls for the average female labor participation rate in the respondents' country of birth. This variable has been the most widely used proxy to capture labor-related gender norms in an immigrant's country of birth by the epidemiological approach to culture (Fernández and Fogli 2009; Fernández 2011; Blau and Kahn 2015).¹³ To ensure the consistency of this measure across time and countries, we use the ratio of female to male labor participation rates rather than the raw female labor participation rates. Moreover, we assign the value of this variable at the time of immigration of the respondent to better capture the conditions in which she formed her preferences regarding gender roles. We also control for the average labor participation rate of married female immigrants to the U.S. from the respondent's country of birth a decade before the time of her migration. We construct this variable using the U.S. censuses from 1940 to 2000 and the ACS from 2010 to 2014 (Ruggles et al. 2015). This approach allows us to address potential forms of selection regarding the culture of the country of origin of respondents that may be different from the culture of the average citizen, since immigrants are a selected pool. It also allows us to capture some degree of oblique cultural transmission by looking at whether the behavior of same country immigrants that arrived to the U.S. a decade earlier than the respondent could play an independent role. To account for the influence of historical contact across populations and the development of gender norms among groups, we also control for a measure of genetic distance from the U.S.

¹² Indeed, as can be seen in the last column of Table 1, Asian immigrants are more likely to speak a non sex-based language. Since they have on average higher levels of labor force participation than other respondents, this explains part of the decline in magnitude.

¹³ To capture cultural variation in gender roles, existing studies have proxied culture with female outcomes in the country of origin. For instance, Fernández and Fogli (2009) use for country of origin female labor force participation to capture the culture of second generation immigrants to the U.S. Blau and Kahn (2015) additionally control for individual labor force participation prior to migration to separate culture from social capital. Oreffice (2014) create an index of gender roles in the country of origin as a function of several gender outcomes. This literature posits that immigrants carry with them some of the attitudes from their home country to the U.S., and, in the case of second generation immigrants, that immigrants transmit some of these attitudes to their children.

(Spolaore and Wacziarg 2009, 2016).¹⁴ We further control for various country-level characteristics such as GDP per capita, total fertility rate, an indicator for whether the country of birth shares a common language with those spoken by at least 9% of the population in the U.S.—English and Spanish—, years of schooling in the country of birth, and various geographic measures—latitude, longitude, and bilateral distance to the U.S.¹⁵ It is important to include these factors since they can control for some omitted variables that correlate with country of origin gender norms.

We find that a one standard deviation increase in female labor force participation rates in one's country of birth (18%) is associated with a 4 percentage points increase in one's labor force participation. Our results imply a larger magnitude for the correlation with the labor force participation of married female immigrants that migrated a decade prior to the respondent, wherein a one standard deviation increase in their average labor participation rate (12%) is associated with a 5 percentage points increase in labor force participation. These results largely confirm previous findings in the literature using the epidemiological approach to culture (Fernández and Fogli 2009; Blau and Kahn 2015).

In column (4), we add the SB measure to compare the magnitude of its impact on individual behavior to that of the usual country-level proxies used in the literature. Including these variables altogether imposes a very stringent test on the data. To see this, refer again to the bottom of Table 3 which reports the residual variation in SB used for identification. This metric is cut in half as we move from column (2) to column (4), implying a correlation between language structure and the usual country-level proxies of gender roles used in the literature.¹⁶ In spite of this, while the coefficient on sex-based language diminishes somewhat in magnitude, it remains highly statistically significant, sizeable, and negative even after adding these controls. This suggests that language structure may capture unobservable cultural characteristics at the country-level beyond the proxies used in the literature. This is promising as it shows that language likely captures some cultural features not previously uncovered.

Nevertheless, while column (4) controls for an exhaustive list of country of birth characteristics as well as language structure, there may still be some unobservable cultural components that vary systematically across countries and that can be captured neither by language nor by country of birth controls. Since immigrants from the same country may speak languages with varying structures, our analysis can uniquely address this issue by further including a set of country of birth fixed effects

¹⁴ Alternatively, we checked the robustness of our results to measures of linguistic distance between languages from Adsera and Pytlikova (2015), including a Linguistic proximity index constructed using data from *Ethnologue*, and a Levenshtein distance measure developed by the *Max Planck Institute for Evolutionary Anthropology*. This data covers 42 languages out of the 63 in our regression sample. Even with a reduced sample size (435,899 observations instead of 480,619 observations), the results from the baseline regression the regression from Table 3 column (5) with country of birth fixed effects are essentially unchanged. Because of such a lower coverage of languages the use of linguistic distance as a control would entail, we do not include these measures throughout the analysis.

¹⁵ See Online Appendix A.2.6 for more details about the sources and the construction of the country-level variables. See also Online Appendix Table C.5, which reports summary statistics for these variables.

¹⁶ This is not surprising given the cross-country results in Gay et al. (2013), which show that country-level female labor participation rates are correlated with the linguistic structure of the majority language.

and exploiting within country variation in the structure of language spoken in the pool of immigrants in the U.S. Because these fixed effects absorb the impact of all time invariant factors at the country level, this means that identification of any language effect relies on heterogeneity in the structure of language spoken within a pool of immigrants from the same country of origin. The addition of 134 country of birth fixed effects noticeably reduce the potential sources of identification, as the residual variance in **SB** is only one tenth of its original value. Yet again, the impact of language remains highly significant and economically meaningful. Gender assigned females are 2.7 percentage points less likely to participate in the labor force than their non-gender assigned counterparts. This suggests that while the cultural components common to all immigrants from the same origin country carried in the structure of language drive a large part the results compared to the estimates in column (2), about one third of the effect can still be attributed to either more local components of culture or to alternative channels such as a cognitive mechanism through which language would impact behavioral outcomes directly.

Note that because identification now comes from within country variation in the structure of languages spoken, the estimates, while gaining in credibility, could be decreasing in representativeness. For instance, they no longer provide information about the impact of language structure on the working behavior of female immigrants that are from linguistically homogeneous countries. To better understand the extent to which each country of birth contributes to the identification, we adapt Aronow and Samii's (2016) procedure to uncover the "effective sample" used in the regression. This procedure generates regression weights by computing the relative size of the residual variance in the **SB** variable for each country of birth in the sample. Not surprisingly, we find that immigrants born in linguistically heterogeneous countries are the prime contributors in building the estimate. In fact, empirical identification in the full fixed effect regressions comes from counties such as India, the Philippines, Vietnam, China, Afghanistan, and Canada. Online Appendix Figure D.1 makes this point clearer by mapping these regression weights.

Despite the fact that culture is a slow-moving institution, it can evolve, and language may itself constrain or facilitate cultural evolution in certain directions. We check that this does not impact our results by adding country of birth fixed effects interacted with decade of migration in column (6). This allows us to capture, to some extent, country-level cultural aspects that could be time variant. The results are largely unchanged by this addition, suggesting that language structure captures permanent aspects that operate at the individual level.

Finally, note that all the regressions reported in Table 3 include state of residence fixed effects to account for the possibility that location choices are endogenous to the language spoken by the community of immigrants that the respondent belongs to. However, this phenomenon could operate at a lower geographic level than that of the state. Therefore, we include county of residence fixed effects in column (7), so that we can effectively compare female immigrants that reside in the same county but speak a language with a varying structure. Unfortunately, the ACS do not systematically provide respondents' county of residence, so that we are only able to include 80% of the respondents in the original regression sample—see Online Appendix A.2.3 for more details. As a results, the regression coefficient in column (7) is not fully comparable with the results in other specifications. Nevertheless, the magnitude and significance of the

estimate on the SB variable remains largely similar to the one in column (6), suggesting that if there is selection into location, it does not drive our main results.

Overall, our findings strongly suggest that while sex-based distinctions in language are deeply rooted in historical cultural forces, gender in language appears to retain a distinct association with gender in behavior which is independent of these other factors.

3.2 Language and the household

Given the failure of the unitary model of household decision-making, standard theory has developed frameworks where bargaining is key to explaining household behavior (Chiappori et al. 2002; Blundell et al. 2007). In this section we analyze whether the impact of language is mediated by household characteristics in the following ways. First, household bargaining power may mediate or influence the impact of language. For example, females with high bargaining power may not be bound by the gender roles that languages can embody. To shed light on the mechanism behind the association between language and female labor supply, we analyze in Section 3.2.1 the impact of the distribution of bargaining power within the household. Second, we investigate in Section 3.2.2 the impact of the language spoken by the husband. While the majority of immigrant households in the data are linguistically homogeneous, roughly 20% are not. This variation allows us to analyze the relative role of the language structures of both spouses, potentially shedding additional light on whether language use within the household matters. Since marriage is more likely among individuals who share the same language, we need to rule out the possibility that our results overestimate the impact of language via selection effects. This could be the case if marriages into linguistically homogeneous languages reflected attachment to one's own culture. To deal with such selection issues, we compare the behavior of these two types of households and exploit information on whether marriage predates migration.

3.2.1 Language and household bargaining power

In this section, we examine the impact of language taking into account bargaining power characteristics within the household. Throughout, we restrict the sample to households where both spouses speak the same language. We consider these households to avoid conflating other potential language effects. This exercise complements recent theoretical advances by analyzing whether the impact of bargaining power in the household is culturally dependent. In particular, the collective model of labor supply predicts that women with lower bargaining power work more, while those with higher bargaining power work less since they are able to substitute leisure for work. At the same time, these model typically do not consider social norms. This may be problematic, as Field et al. (2016) show that including social norms into a collective labor supply model can lead to the opposite prediction, namely that women with lower bargaining power work less.¹⁷

¹⁷ Field et al. (2016) implement a field experiment in India where traditional gender norms bound women away from the labor market, and investigate how a change in their bargaining power stemming from an increase in their control over their earnings can allow them to free themselves from the traditional gender norms.

For the sake of comparison, we report in column (1) of Table 4 the results when replicating column (5) of Table 3 with this alternative sample of linguistically homogeneous households. Women speaking a sex-based language are 3.1 percentage points less likely to participate in the labor force than their non-gender assigned counterparts. This is slightly larger in magnitude than in the unrestricted sample, where the effect is of 2.7 percentage points.

To capture bargaining power within the household, we follow Oreffice (2014) and control for the age gap as well as the non-labor income gap between spouses.¹⁸ In column (2), we exclude the language variable to derive a baseline when using these new controls. These regressions include controls for husbands characteristics similar to those of the respondents used in Table 3.¹⁹ Consistent with previous work, we find that the larger the age gap and the larger the non-labor income gap, the less likely women are to participate in the labor force: a one standard deviation increase in the age gap is associated with a reduction in the wives labor force participation of 1.8 percentage point. Similarly, a one standard deviation increase in the non-labor income gap is associated with a reduction in the wives labor force participation of 2.1 percentage points. These results corroborate Oreffice (2014) and Field et al. (2016) which show that the impact of bargaining power is different in households with traditional gender roles relative to the U.S., and that female immigrants with low bargaining power are less likely to participate in the labor market.²⁰

In column (3), we add the SB variable. The coefficient is largely unchanged compared to column (1) as women speaking a sex-based language are 2.7 percentage points less likely to participate in the labor force than their non-gender assigned counterparts. This suggests that language has an impact that is not mediated by the distribution of bargaining power within the household. Furthermore, the impact of language is of comparable magnitude as that of either the non-labor income gap or the age gap, suggesting that cultural forces are equally important than bargaining measures in determining female labor forces participation.

In 16% of households where both spouses speak the same language, spouses were born in different countries. To account for any impact this may have on the estimates, we add husband country of birth fixed effects in column (4). Reassuringly, the results are largely unchanged by this addition. Similarly, some of the observed effect may result from selection into same culture marriages. To assess whether such selection effect drives the results, we run the specification of column (4) on the subsample of spouses that married before migration. Since female immigrants married to their husbands prior to migration (“tied women”) may have different motivations, this subsample should provide a window into whether we should worry about selection

¹⁸ We focus on the non-labor income gap because it is relatively less endogenous than the labor income gap (Lundberg et al. 1997). We do not include the education gap for the same reason. Online Appendix Table C.4 presents descriptive statistics for various gender gap measures within the household. Other variables, such as physical attributes, have been shown to influence female labor supply (Oreffice and Quintana-Domeque 2012). Unfortunately, they are not available in the ACS (Ruggles et al. 2015).

¹⁹ Online Appendix Table C.3 presents descriptive statistics for these husband characteristics.

²⁰ This also consistent with findings in Hicks et al. (2015), wherein the division of household labor was shown to be heavily skewed against females in households coming from countries with a dominantly sex-based language, suggesting that languages could potentially constrain females to a traditional role within the household.

Table 4 Language structure and household bargaining: Same Language married couples (2005–2015)

Dependent variable:	Labor force participation					
	(1)	(2)	(3)	(4)	(5)	(6)
Sex-based	-0.031*** [0.008]		-0.027*** [0.008]	-0.022** [0.009]	-0.021* [0.013]	-0.023*** [0.009]
β -coef.	-0.031		-0.027	-0.022	-0.021	-0.022
Age gap		-0.003*** [0.001]	-0.003*** [0.001]	-0.006*** [0.001]	0.003 [0.002]	-0.007*** [0.001]
β -coef.		-0.018	-0.018	-0.034	0.013	-0.035
Non-labor income gap		-0.001*** [0.000]	-0.001*** [0.000]	-0.001*** [0.000]	-0.001*** [0.000]	-0.001*** [0.000]
β -coef.		-0.021	-0.021	-0.021	-0.018	-0.015
Age gap \times SB						0.000 [0.000]
β -coef.						0.002
Non-labor income gap \times SB						-0.000*** [0.000]
β -coef.						-0.008
Respondent char.	Yes	Yes	Yes	Yes	Yes	Yes
Household char.	Yes	Yes	Yes	Yes	Yes	Yes
Respondent COB FE	Yes	Yes	Yes	Yes	Yes	Yes
Husband char.	No	Yes	Yes	Yes	Yes	Yes

Table 4 continued

Dependent variable:	Labor force participation					
	(1)	(2)	(3)	(4)	(5)	(6)
Husband COB FE	No	No	No	Yes	Yes	Yes
Sample mar. before mig.	No	No	No	No	Yes	No
Observations	409,017	409,017	409,017	409,017	154,983	409,017
R ²	0.134	0.145	0.145	0.147	0.162	0.147
Mean	0.59	0.59	0.59	0.59	0.55	0.59
SB residual variance	0.011		0.011	0.010	0.012	0.013

Notes: The estimates are computed using sample weights (PERWT) provided in the ACS (Ruggles et al. 2015). See Online Appendix A for details on variable sources and definitions

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level

among couples after migration. Note that because this information is only available in the ACS after 2008, the resulting estimate is not fully comparable to others in Table 4. Nevertheless, the magnitude of the estimate on the language variable remains largely unchanged—although the dramatic reduction in the sample size reduces our statistical power. The impact of non-labor income gap also remains roughly similar, while the age gap coefficient becomes insignificant and positive. Overall, these results suggest that our main findings are not driven by selection into linguistically homogeneous marriages.

Again motivated by Orefice (2014) and Field et al. (2016), we investigate the extent to which gender roles, as embodied by gender in language, are reinforced in households where the wife has weak bargaining power or vice versa. In column (6) of Table 4, we add interaction terms between the SB language variable and the non-labor income gap as well as the age gap. After this addition, the estimate on the language measure remains identical to the one in column (4), suggesting that language has a direct effect that is not completely mediated through bargaining power. The only significant interaction is between the language variable and the non-labor income gap. In particular, a one standard deviation increase in the size of the non-labor income gap leads to an additional decrease of 0.8 percentage points in the labor participation of women speaking a sex-based language compared to others that does not. While the effect is arguably small in magnitude, it does confirm Field et al. (2016) insofar as married females with low bargaining power are more likely to be bound by traditional gender roles, and less likely to participate in the labor market. Online Appendix Table C.12 replicates the analysis carried out in column (6) alternative measures of labor market engagement. As in previous analyses, the results are similar for the extensive margin but less clear for the intensive margin.

3.2.2 Evidence from linguistic heterogeneity within the household

So far, we have shown that the impact of gender in language on gender norms regarding labor market participation was robust to controlling for husband characteristics as well as husband country of origin fixed effects. In this section, we analyze the role of gender norms embodied in a husband's language on his wife's labor participation. Indeed, Fernández and Fogli (2009) find that gender roles in the husband's country of origin characteristics play an important role in determining the working behavior of his wife.

In Table 5, we pool together all households, and compare households where both spouses speak a sex-based language to households where only one speaker does so, paying attention to whether this speaker is the husband or the wife.²¹ We also consider husbands that speak English. Because English speaking husbands may provide assimilation services to their wives, we include an indicator for English

²¹ Throughout the table we sequentially add female country of birth characteristics and female country of birth fixed effects. We do not add the respondent's husband country of birth fixed characteristics because only 20% of the sample features husbands and wives speaking different structure languages. In that case, our analysis does not have sufficient statistical power precisely identify the coefficients. While many non-English speaking couples share the same language in the household, we focus on the role of gender marking to learn about the impact of husbands and wives gender norms as embodied in the structure of their language.

Table 5 Linguistic heterogeneity in the household immigrant households (2005–2015)

Dependent variable	Labor force participation				
	(1)	(2)	(3)	(4)	(5)
Same × wife's sex-based	-0.068*** [0.003]	-0.069*** [0.003]	-0.070*** [0.003]	-0.040*** [0.005]	-0.015* [0.008]
Different × wife's sex-based		-0.043*** [0.014]	-0.044*** [0.014]	-0.015 [0.015]	0.006 [0.016]
Different × husband's sex-based		-0.031*** [0.011]	-0.032*** [0.011]	-0.008 [0.012]	0.001 [0.011]
Respondent char.	Yes	Yes	Yes	Yes	Yes
Household char.	Yes	Yes	Yes	Yes	Yes
Husband char.	Yes	Yes	Yes	Yes	Yes
Respondent COB char.	No	No	No	Yes	No
Respondent COB FE	No	No	No	No	Yes
Observations	387,037	387,037	387,037	364,383	387,037
R ²	0.122	0.122	0.122	0.143	0.146
Mean	0.59	0.59	0.59	0.59	0.59
SB residual variance	0.095	0.095	0.095	0.043	0.012

Notes: The estimates are computed using sample weights (perwt) provided in the ACS (Ruggles et al. 2015). See Online Appendix A for details on variable sources and definitions

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level

speaking husbands to sort out the impact of the language structure from these assimilation services.

In column (1) of Table 5, the wife's language variable **SB** is interacted with two indicators that capture whether a wife and her spouse both speak a language that has the same grammatical structure.²² Women speaking a sex-based language in a household with a husband speaking a sex-based language as well are almost 7 percentage points less likely to be in the labor force. Conversely, women speaking a sex-based language in a household with a husband speaking a non sex-based language are only 4.3 percentage points less likely to be part of the labor force. To further explore the role of the husbands' languages, we run in column (2) the same specification as in column (1) except that we use the husband's language variable rather than the wife's. The results are very close to those in column (1) but suggest that while a husband's gender norms matter for a wife's behavior, they seem to play a slightly weaker albeit still significant role. Column (3) includes both spouses'

²² Instead of simply comparing households where husbands and wives speak the same language vs. those without, this approach allows us both to include households with English speakers and to understand whether the observed mechanisms operate through grammatical gender.

language variables. Overall, this suggests that the impact of language structure is stronger when the languages of both spouses are sex-based. Furthermore, when speaking languages with different gender structures, the impact of the respondent's language is bigger than the impact of the husband's language, suggesting cultural spillovers within the household. Finally, columns (4) and (5) of Table 5 add country of birth characteristics and country of birth fixed effects, respectively. The impact of the husband's language remains significant in linguistically homogeneous households, but not in heterogeneous ones. Nevertheless, it seems that part of this loss in precision is due to the steep decline in statistical power, as testified by the decline in the residual variance in SB.²³

3.3 Language and social interactions

Immigrants tend to cluster in ethnic enclaves, partly because doing so allows them to access a network where information is exchanged (Edin et al. 2003; Munshi 2003). Within those networks, speaking a common language may facilitate such exchanges. Furthermore, language itself is a network technology which value increases with the number of speakers. Being able to communicate within a dense ethnic network may be particularly important for female immigrants, to share information, communicate about job opportunities, and reduce information asymmetries between job seekers and potential employers. This effect may encourage female labor force participation. At the same time, sharing the same linguistic and cultural background within a dense ethnic network may reinforce the social norms that the language act as a vehicle for. This is even more so if increasing language use makes gender categories more salient, or if sharing the same cultural background makes social norms bind more strongly. As a result, female immigrants living in an ethnic enclave may face a trade-off between, on the one hand, increased job opportunities through informal network channels, and, on the other hand, increased peer pressure to conform with social norms. This second effect may itself depend on the extent to which the ethnic group social norms are biased in favor of traditional gender roles that encourage women to stay out of the labor force.

This potential trade-off guides the empirical analysis presented in Table 6. In what follows, we rely on the subsample of immigrants residing in the counties that are identifiable in the ACS. This corresponds to roughly 80% of the initial sample.²⁴ To capture the impact of social interactions, we build two measures of local ethnic and linguistic network density. *Density COB* is the ratio of the immigrant population born in the respondent's country of birth and residing in the respondent's county to the total number of immigrants residing in the respondent's county. Similarly, *Density language* is the ratio of the immigrant population speaking the respondent's

²³ In all cases however, when a wife speaks a sex-based language, she exhibits on average lower female labor force participation.

²⁴ See Online Appendix A.2.3 for more details on the number of identifiable counties in the ACS. Although estimates on this subsample may not be comparable with those obtained with the full sample, the results column (7) of Table 3 being so similar to those in column (6) gives us confidence that selection into county does not drive the results.

Table 6 Language and social interactions respondents in identifiable counties (2005–2015)

Dependent variable:	Labor force participation					
	(1)	(2)	(3)	(4)	(5)	(6)
Sex-based		-0.023*** [0.003]		-0.037*** [0.003]	-0.026*** [0.003]	-0.016* [0.008]
β -coef.		-0.225		-0.245	-0.197	-0.057
Density COB	-0.001*** [0.000]	0.009*** [0.000]			0.009*** [0.001]	0.003*** [0.001]
β -coef.	-0.020	0.223		0.221		0.067
Density COB \times SB		-0.009*** [0.000]			-0.010*** [0.001]	-0.002*** [0.001]
β -coef.		-0.243		-0.254		-0.046
Density language			0.000** [0.000]	0.007*** [0.000]	-0.000 [0.001]	-0.000 [0.001]
β -coef.			0.006	0.203	-0.001	-0.007
Density language \times SB				-0.007*** [0.000]	0.001** [0.001]	-0.000 [0.001]
β -coef.				-0.201	0.038	-0.004
Respondent char.	Yes	Yes	Yes	Yes	Yes	Yes
Household char.	Yes	Yes	Yes	Yes	Yes	Yes
Respondent COB FE	No	No	No	No	No	Yes
Observations	382,556	382,556	382,556	382,556	382,556	382,556
R ²	0.110	0.113	0.110	0.112	0.114	0.134
Mean	0.60	0.60	0.60	0.60	0.60	0.60
SB residual variance		0.101		0.101	0.101	0.013

Notes: The estimates are computed using sample weights (PEFRT) provided in the ACS (Ruggles et al. 2015). See Online Appendix A for details on variable sources and definitions

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level

Table 7 Gender in language and economic participation, additional outcomes replication of column 5 of Table 3, all language variables

Dependent variable:	Extensive margin		Intensive margins			
	LFP (1)	Employed (2)	Including zeros		Excluding zeros	
			Weeks (3)	Hours (4)	Weeks (5)	Hours (6)
Panel A. Sex-based						
SB	-0.027*** [0.007]	-0.028*** [0.007]	-1.01*** [0.34]	-0.813*** [0.297]	-0.24 [0.24]	-0.165 [0.229]
Observations	480,618	480,618	480,618	480,618	302,653	302,653
R ²	0.132	0.135	0.153	0.138	0.056	0.034
Panel B. Number of genders						
NG	-0.018*** [0.005]	-0.023*** [0.005]	-1.00*** [0.22]	-0.708*** [0.193]	-0.51*** [0.16]	-0.252* [0.136]
Observations	480,618	480,618	480,618	480,618	302,653	302,653
R ²	0.132	0.135	0.154	0.138	0.056	0.034
Panel C. Gender assignment						
GA	-0.011** [0.005]	-0.017*** [0.005]	-0.69*** [0.25]	-0.468** [0.219]	-0.51*** [0.18]	-0.335** [0.155]
Observations	480,618	480,618	480,618	480,618	302,653	302,653
R ²	0.132	0.135	0.153	0.138	0.056	0.034

Table 7 continued

Dependent variable:	Extensive margin		Intensive margins			
	LFP (1)	Employed (2)	Including zeros		Excluding zeros	
			Weeks (3)	Hours (4)	Weeks (5)	Hours (6)
Panel D. Gender pronouns						
GP	-0.019** [0.009]	-0.013 [0.009]	-0.44 [0.42]	-0.965*** [0.353]	0.12 [0.31]	-0.849*** [0.275]
Observations	478,883	478,883	478,883	478,883	301,386	301,386
R ²	0.132	0.135	0.153	0.138	0.056	0.034
Panel E. Intensity						
Intensity	-0.010*** [0.003]	-0.012*** [0.003]	-0.51*** [0.12]	-0.416*** [0.105]	-0.26*** [0.09]	-0.222*** [0.076]
Observations	478,883	478,883	478,883	478,883	301,386	301,386
R ²	0.132	0.135	0.153	0.138	0.056	0.034
Respondent char.	Yes	Yes	Yes	Yes	Yes	Yes
Household char.	Yes	Yes	Yes	Yes	Yes	Yes
Respondent COB FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.60	0.55	27.2	22.51	44.2	36.57

Notes: The estimates are computed using sample weights (PERWT) provided in the ACS (Ruggles et al. 2015). See Appendix A for details on variable sources and definitions
 ***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level

language and residing in the respondent's county to the total number of immigrants residing in the respondent's county.²⁵

Throughout Table 6, we control for respondent characteristics and household characteristics. In the full specification we also control for respondent country of birth fixed effects. In column (1) of Table 6, we exclude the language variable and include the density of the respondent's country of birth network. The density of the network has a negative impact on the respondent's labor participation, suggesting that peer pressure from immigrants coming from the same country of origin to comply with social norms may be stronger than improvements in access to job opportunities. In column (2), we add both the language variable and the interaction term between language and the network density measure. The magnitude of the coefficients imply that language and network densities play different roles that reinforce each other. In particular, the results suggest that peer pressure to comply with social norms may be stronger among female immigrants that speak a sex-based language, as the interaction term is negative. Also, the impact of the respondent's network is now strongly positive, suggesting that, on its own, living in a county with a dense ethnic network does increase labor force participation once cultural factors are taken into account. Finally, the coefficient on the language variable is still negative and significant, and has the same magnitude as in all other specifications. In columns (3) and (4), we repeat the same exercise using instead the network density in language spoken as criteria to define ethnic density. The results and interpretation are broadly similar to the previous ones. Finally, we add both types of network density measures together in columns (5) and (6). The results suggest that ethnic networks defined using country of birth matter more than networks defined purely along linguistic lines.

3.4 Robustness checks

3.4.1 Alternative language variables and labor market outcomes

This section presents an extensive list of robustness checks of the main results in Table 3, column (5)—a specification that includes country of birth fixed effects. First, Table 7 reports the results when using alternative measures of female labor participation, together with alternative language variables.

Columns (1) and (2) of Table 7 show the results for two measures of the extensive margin: labor force participation and an indicator for being employed, respectively. We obtain quantitatively and statistically comparable results for both measures when using our main language variable, SB, which confirms our previous analysis. Regarding the intensive margin, we also use the number of weeks worked and hours worked as dependent variables, including zeros in columns (3) and (4), and excluding zeros in columns (5) and (6). This provides various measures for whether language influences not only whether women work, but also how much. The impact of language is stronger for the extensive margin than for the intensive margin. For instance, although the impact of language is not precisely estimated when the

²⁵ In both measures, we use the total number of immigrants that are in the workforce because it is more relevant for networking and reducing information asymmetries regarding labor market opportunities. See Online Appendix A.2.7 for more details.

dependent variable is the number of weeks worked excluding zeros in column (5), the magnitude is very small as female immigrants with a gender-based language work on average one and a half day less per year. This corresponds to only half a percent of the sample mean. This suggests that language acts as a vehicle for traditional gender roles that tend to ascribe women to the household, and to exclude them from the labor market. Once they overcome such roles by participating in the labor force, the impact of language remains, although it is weaker. This is potentially due to gender norms that unevenly distribute the burden of household tasks even among couples where both partners work, decreasing the labor supply of those female workers (Hicks et al. 2015).

In panels B, C, and D of Table 7, we replicate our analysis of panel A with each of the individual measures of gender marking discussed in Section 2.1. In all cases, we obtain consistent results: married female immigrants speaking a language with gender distinctions are less likely to work, and conditional on working, they are doing so less intensively—although the magnitude of the impact of language on the intensive margin is smaller than that on the extensive margin. Finally, panel E reports the results when using the composite index Intensity described in Section 2.1. The results in column (1) show that in comparison to those speaking a gender marked language with the lowest intensity (Intensity = 0), female immigrants speaking a language with the highest gender intensity (Intensity = 3) are 3 percentage points less likely to be in the labor force. The results are similar when using alternative outcome variables. Moreover, the estimates are more precisely estimated than with the indicator variables for language structures. In Online Appendix Table C.7, we show that the results in panel E are not sensitive to the specification of the Intensity measure, as the results hold with four alternative specifications of the index.

3.4.2 Alternative samples

Online Appendix Table C.8 explores the robustness of our main results from column (5) of Table 3 to alternative samples.²⁶ We use a wider age window (15–59) for the sample in column (2). Results are similar to the baseline, suggesting that education and retirement decisions are not impacted by language in a systematic direction. In column (3), we check that identification is not driven by peculiar migrants by restricting the sample to respondents speaking a language that is indigenous to their country of birth, where we define a language as not indigenous to a country if it is not listed as a principal language spoken in a country in the *Encyclopedia Britannica Book of the Year* (2010, pp. 766–770). We also checked that the results are not driven by outliers, and robust to excluding respondents from countries with less than 100 observations—this is the case for 1064 respondents—and respondents speaking a language that is spoken by less than 100 observations—this is the case for 86 respondents. The results are again similar. We run the baseline specification on other subsamples as well: we include English speaking immigrants in column (4), exclude Mexican immigrants in column (5), include all types of households in column (6), and exclude languages that have been imputed as indicated by the quality flag

²⁶ See Online Appendix A.1.2 for more details on how we constructed these subsamples.

QULANGUAG in the ACS (Ruggles et al. 2015). The results are robust to these alternative samples.

3.4.3 *Alternative functional forms*

We also undertook robustness checks concerning the empirical specification. In particular, we replicate in Online Appendix Table C.9 the main results in columns (1) and (5) of Table 3 using both a probit and logit model. The marginal estimates evaluated at the mean are remarkably consistent with the estimates obtained via the OLS linear probability models. We take this as evidence that the functional form is not critical.²⁷

3.4.4 *Respondents husbands' labor supply*

Another important robustness check concerns the impact of language on the labor supply of the respondents' husbands. It is important to rule out the possibility that we observe the same effects than for the female respondents, namely that sex-based language speaking husbands are less likely to engage in the labor force as well. This would indicate that our results are spurious and unrelated to traditional gender norms. To verify this, Online Appendix Table C.10 replicates columns (1), (2), and (5) of Table 3, with the labor supply of the respondents' husbands as the dependent variable, and with their characteristics as controls. So that the sample is qualitatively similar to the one used in the baseline analysis, we exclude native husbands as well as English speaking husbands. We find a significant positive association between husbands' SB language variable and their labor force participation for the specifications without husband country of birth fixed effects. This suggests that our language variable captures traditional gender roles in that it leads couples to a traditional division of labor where wives stay home and husbands work. Yet, once we control for husband country of birth fixed effects, the association is no longer statistically significant, suggesting that the influence of this cultural trait is larger for women. Overall this analysis reassures us that our results are not driven by some correlated factor, which would lead speakers of sex-based language to decrease their labor market engagement regardless of their own gender.

3.4.5 *Heterogeneity across marital status*

Finally, we also explore potential heterogeneity in the impact of language structure across marital status. Although married women represent 83% of the original uncorrected sample, it is worth analyzing whether the impact of language is similar for unmarried women. In Online Appendix Table C.11, we replicate column (5) of Table 3 with an additional indicator variable for unmarried respondents, as well as the interaction of this indicator with the SB variable. Depending on the specification, it reveals that unmarried women are 8 to 14 percentage points more likely to participate in the labor force compared to married women. Second, sex-based speakers

²⁷ We maintain the OLS throughout the paper, however, as it is computationally too intensive to run these models with the inclusion of hundreds of fixed effects in most of our specifications.

are less likely than their counterparts to be in the labor force when they are married, but the reverse is true when they are unmarried: while married women speaking a sex-based language are 5 percentage points less likely to be in the labor force than their counterparts, unmarried women speaking a sex-based language are 8 percentage points *more* likely to work than their counterparts. When paired with the findings for single women in Hicks et al. (2015), this result suggests that some forms of gender roles may be “dormant” when unmarried—the pressure to not work, to raise children, to provide household goods—and that these forces may activate for married women but not be present for unmarried women. Other gender norms and choices, such as deciding how much time to devote to household chores such as cleaning, may be established earlier in life and may appear even in unmarried households Hicks et al. (2015).

4 Conclusion

This paper contributes to the existing literature on the relationship between grammatical features of language and economic behavior by examining the behavior of immigrants, who travel with acquired cultural baggage, including their language. While no quasi-experimental study is likely to rule out all potential sources of endogeneity, our data driven, fixed effects, epidemiological analysis advances the existing frontier in the economic analysis of language and provides suggestive evidence that the study of language deserves further attention.

We provide support for the nascent strand of literature in which languages serve not only to reflect, but also possibly to reinforce and transmit culture. In particular, our quantitative exercise isolates the fraction of this association attributable to gendered language from the portion associated with other cultural and gender norms correlated with language. We find that about two thirds of the correlation between language and labor market outcomes can be attributed to the latter, while at most one third can be attributed to the direct impact of language structure or other time-variant cultural forces not captured by traditional observables or by the wide array of additional checks we employ. Whether by altering preference formation or by perpetuating inefficient social norms, language and other social constructs clearly have the potential to hinder economic development and stymie progress of gender equality.

We frame our analysis within a collective household labor supply model and demonstrate that language has a direct effect that is not strongly influenced by either husband characteristics or the distribution of bargaining power within the household. This suggests that language, and more broadly acquired gender norms, should be considered in their own right in analyses of female labor force participation. In this regard, language is especially promising since it allows researchers to study a cultural trait which is observable, quantifiable, and varies at the individual level.

Furthermore, we show that the labor market associations with language are larger in magnitude than some factors traditionally considered to capture bargaining power, in line with Oreffice (2014) who finds that culture can mediate the relationship between bargaining power and the labor supply. Indeed, our findings regarding the impact of language in linguistically homogeneous and heterogeneous households

suggests that the impact of one's own language is the most robust predictor of behavior, although the spouse's language is also associated with a partner's decision making.

Finally, recognizing that language is a network technology allows us to examine the role that ethnic enclaves play in influencing female labor force participation. In theory, enclaves may improve labor market outcomes by providing information about formal jobs and reducing social stigma on employment. At the same time, enclaves are likely to provide isolation from U.S. norms and to reinforce gender norms that languages capture, enhancing the impact of gender in language. We present evidence suggesting that the latter effect is present. Explaining the role of language within social networks, therefore may shed new light on how networks may impose not only benefits but also costs on its members by reinforcing cultural norms.

Our results have important implications for policy. Specifically, programs designed to promote female labor force participation and immigrant assimilation could be more appropriately designed and targeted by recognizing the existence of stronger gender norms among subsets of speakers. Future research may consider experimental approaches to further analyze the impact of language on behavior, and in particular, to better understand the policy implications of movements for gender neutrality in language. Another interesting avenue for research might be to study the impact of gendered grammatical features in a marriage market framework as in Grossbard (2015), which studies intermarriage among immigrants along linguistic lines.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

- Adsera, A., & Pytlikova, M. (2015). The role of language in shaping international migration. *The Economic Journal*, 125(586), F49–F81.
- Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2), 469–530.
- Aronow, P. M., & Samii, C. (2016). Does regression produce representative estimates of causal effects? *American Journal of Political Science*, 60(1), 250–267.
- Blau, F. D., & Kahn, L. M. (2015). Substitution between individual and source country characteristics: Social capital, culture, and us labor market outcomes among immigrant women. *Journal of Human Capital*, 9(4), 439–482.
- Blau, F., Kahn, L., & Papps, K. (2011). Gender, source country characteristics, and labor market assimilation among immigrants. *The Review of Economics and Statistics*, 93(1), 43–58.
- Blundell, R., Chiappori, P.-A., Magnac, T., & Meghir, C. (2007). Collective labour supply: Heterogeneity and non-participation. *The Review of Economic Studies*, 74(2), 417–445.
- Chen, M. K. (2013). The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *American Economic Review*, 103(2), 690–731.
- Chiappori, P.-A., Fortin, B., & Lacroix, G. (2002). Marriage market, divorce legislation, and household labor supply. *Journal of Political Economy*, 110(1), 37–72.
- Davis, L., & Reynolds, M. (2016). Gendered language and the educational gender gap. Available at SSRN: <https://ssrn.com/abstract=2782540>.

- Dryer, M., & Haspelmath, M. (2011). The world atlas of language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info>).
- Edin, P.-A., Fredriksson, P., & Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants: evidence from a natural experiment. *The Quarterly Journal of Economics*, 118(1), 329–357.
- Everett, C. (2013). *Linguistic relativity: Evidence across languages and cognitive domains*, Vol. 25, Berlin/Boston: Walter de Gruyter GmbH.
- Farré, L., & Vella, F. (2013). The intergenerational transmission of gender role attitudes and its implications for female labour force participation. *Economica*, 80(318), 219–247.
- Fernández, R. (2007). Alfred marshall lecture women, work, and culture. *Journal of the European Economic Association*, 5(2–3), 305–332.
- Fernández, R. (2011). Chapter 11 - does culture matter?, Vol. 1 of *Handbook of social economics* (pp. 481–510). North Holland: Elsevier.
- Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *The American Economic Review*, 103(1), 472–500.
- Fernández, R., & Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1), 146–177.
- Field, E., Pande, R., Rigol, N., Schaner, S., & Moore, C. T. (2016). On her account: Can strengthening women's financial control boost female labor supply?, Technical report.
- Gay, V., Hicks, D., & Santacreu-Vasut, E. (2016). Migration as a window into the coevolution between language and behavior. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (EVOLANGX11)*, Online at <http://evolang.org/neworleans/papers/120.html>.
- Gay V., Santacreu-Vasut E., & Shoham A. (2013). *The grammatical origins of gender roles* (Working Paper Series WP2013-03). Berkeley, CA: Berkeley Economic History Economic Laboratory.
- Givati, Y., & Troiano, U. (2012). Law, economics, and culture: Theory of mandated benefits and evidence from maternity leave policies. *Journal of Law and Economics*, 55(2), 339–364.
- Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review*, 104(4), 1091–1119.
- Grossbard, S. (2015). *The marriage motive: A price theory of marriage. How marriage markets affect employment, consumption and savings*. Springer-Verlag New York Inc.
- Hicks, D. L., Santacreu-Vasut, E., & Shoham, A. (2015). Does mother tongue make for women's work? linguistics, household labor, and gender identity. *Journal of Economic Behavior & Organization*, 110, 19–44.
- Inc., E. B. (2010). *2010 Britannica book of the year*, Vol. 35, Encyclopaedia britannica. pp. 766–770.
- Ladd, D. R., Roberts, S. G., & Dediu, D. (2015). Correlational studies in typological and historical linguistics. *Annu. Rev. Linguist*, 1(1), 221–241.
- Lundberg, S. J., Pollak, R. A., & Wales, T. J. (1997). Do husbands and wives pool their resources? Evidence from the united kingdom child benefit. *The Journal of Human Resources*, 32(3), 463–480.
- Lupyan G., & Dale R. (2010) Language structure is partly determined by social structure. *PLoS ONE*, 5(1), e8559. doi:10.1371/journal.pone.0008559.
- Mavisakalyan, A. (2015). Gender in language and gender in employment. *Oxford Development Studies*, 43(4), 403–424.
- Mavisakalyan A., & Weber C. (2016) Linguistic Relativity and Economics. Bankwest Curtin Economics Centre Working Paper 16/05, Perth: Curtin University.
- Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, 118(2), 549–599.
- Oreffice, S. (2014). Culture and household decision making. balance of power and labor supply choices of us-born and foreign-born couples. *Journal of Labor Research*, 35(2), 162–184.
- Oreffice, S., & Quintana-Domeque, C. (2012). Fat spouses and hours of work: Are body and pareto weights correlated? *IZA Journal of Labor Economics*, 1(1), 6.
- Roberts S., Winters J., & Chen K. (2015) Future tense and economic decisions: Controlling for cultural evolution. *PLoS ONE*, 10(7), e0132145. doi:10.1371/journal.pone.0132145.
- Roberts S., & Winters J. (2013) Linguistic diversity and traffic accidents: lessons from statistical studies of cultural traits. *PLoS ONE*, 8(8), e70902. doi:10.1371/journal.pone.0070902.
- Ruggles, S., Genadek, K., Goeken, R., Grover, J., & Sobek, M. (2015). *Integrated public use microdata series: Version 6.0*. [dataset]. Minneapolis: University of Minnesota. doi:10.18128/D010.V6.0.
- Santacreu-Vasut, E., Shenkar, O., & Shoham, A. (2014). Linguistic gender marking and its international business ramifications. *Journal of International Business Studies*, 45(9), 1170–1178.

- Santacreu-Vasut, E., Shoham, A., & Gay, V. (2013). Do female/male distinctions in language matter? Evidence from gender political quotas. *Applied Economics Letters*, 20(5), 495–498.
- Spolaore, E., & Wacziarg, R. (2009). The diffusion of development. *The Quarterly Journal of Economics*, 124(2), 469–529.
- Spolaore, E., & Wacziarg, R. (2016). Ancestry and development: New evidence. Discussion Papers Series, Department of Economics, Tufts University 0820, Department of Economics, Tufts University.
- van der Velde, L., Tyrowicz, J., & Siwinska, J. (2015). Language and (the estimates of) the gender wage gap. *Economics Letters*, 136, 165–170.