Intergenerational Transmission of Bilingualism and the Labour Market: Evidence from Mexico

Diego De La Fuente
Department of Economics, University of Sussex

Panu Pelkonen
Department of Economics, University of Sussex
IZA

PRELIMINARY
Please do not cite without authors' permission

30 June 2019

Abstract

This study demonstrates a series of links between minority language skills, their transmission across generations and employment opportunities. Using a detailed matching procedure, we estimate the likelihood of being employed for bilingual versus monolingual men for a large number of Mexican indigenous groups. We find that for indigenous groups, retaining the minority language along with Spanish increases employment opportunities. Furthermore, using the Mexican census household micro data, we show that the languages that are associated with larger labour market benefits, are more likely to be passed on from parents to children, controlling for other factors. Overall, this study shows that continuity of minority languages across generations is linked to concrete economic benefits and labour market specialisation, along with the usual social factors within the family and the community.

Keywords: Intergenerational transmission, language skills, bilingualism, return to skills, minority languages.

1. Introduction

Many developing countries are characterised by a dominant official language (such as English, Spanish or French), and one or many indigenous languages that have varying numbers of speakers or differing levels of official status. In the developed world, minority languages are often beneficiaries of substantial legal protection and educational infrastructure that provides schooling in the minority language. This is typically not the case in developing and emerging economies. In the absence of supportive state institutions, languages either sustain themselves in various social and economic networks, or don't, and face extinction.

In families where at least one parent speaks an indigenous language, parents weigh the options on what languages to teach to children. In a typical case, parents face a choice of raising bilingual children (who speak both the minority and the majority language), or monolingual children (the majority language only). In the common case where the school system supports only the majority language, parents have to make an active effort to maintain the minority language at home.

Current trends in linguistic diversity around the world suggest that minority languages are disappearing fast, and that 90% of the world's languages are expected to disappear in the next 100 years (Nettle and Romaine, 2002). For Mexico, the most important reasons for the erosion of indigenous languages has in the past been a 'forced language shift', an official policy favouring Spanish, but in current times, increasingly also a 'voluntary language shift', abandonment of a language even in the absence of its suppression.

In this study, we examine the transmission of language and the economic rationale of language choice by breaking the family decision to raise bilingual children to different effects. Namely, the strength of social networks and the expected economic return to knowing a minority language. The main social networks are the family and the local area network. As a proxy for the expected economic benefit of bilingualism, we estimate the extent to which bilingual men are more likely to find employment, as opposed to monolingual men.

The study focuses on Mexico, which is one of a handful of large countries which have a rich tapestry of minority languages. Furthermore, the Mexican census may be unique in allowing for a good documentation of the intergenerational transmission of languages within households for a very large sample of people. Another desirable characteristic of the data is that it allows us to distinguish population by both ethnicity and language. The large number of geographically clustered indigenous groups in Mexico creates an exceptional research setting in allowing us to compare the intergenerational transmission rates of languages across these groups, and their varying characteristics and conditions.

Firstly, we use a number of different data sources to show that indigenous men who are bilingual, have on average approximately 2-4 percentage points higher likelihood of being employed, as opposed to observationally similar indigenous men, who can only speak Spanish. These estimates are based on a combination of matching and least squares regression. We also show that the estimated employment return varies by language group, while being positive in most of groups. We present evidence to show this employment effect derives from a larger likelihood of the speakers of indigenous languages to work in agriculture.

In the second part of the study, we rationalise and estimate a model of language transmission within families. We show that only about two-thirds of children with at least one parent who speaks indigenous language, learn to speak it. This proves that a large share of indigenous families in Mexico are in 'the margin' of decision to teach or not to teach the indigenous language to their offspring.

Empirical analysis shows that the key social determinants of language transmission are the number of parents and other adults in the household that can speak the minority language, as well as the share of people in the municipality, who can speak the language. This basic model is further complemented with other parental characteristics such as education level, municipality and regional characteristics.

Once the model is extended to include the estimated employment benefits, which are specific to each indigenous language, we find that larger employment benefits are associated with larger intergenerational transmission of the language, particularly in rural areas. The strong effect in the rural areas is consistent with the fact that a disproportionate share of indigenous people in Mexico live in rural areas and work in traditional occupations such as agriculture or crafts.

Overall, the results suggest that among the numerous indigenous populations of Mexico, knowing the indigenous language allows for broader job opportunities in occupations that the indigenous populations specialise in. As such, learning the indigenous language can be thought of as an 'insurance' against the possibility of an unsuccessful integration to the mainstream job market, where the Spanish language dominates.

The economic literature on language skills is not broad. Particular attention has been paid to the return on language skills of migrants in the developed countries, and the generic, and reasonably well identified conclusion is that immigrants have a high return on fluency in a dominant language (Dustmann 1994, Chiswick and Mille 1995, Dustmann and Fabri 2003, Bleakley and Chin 2004, Miranda and Zhu, 2013). These studies find that immigrants who are proficient in English in the UK, the US or Australia earn 5-36% more, depending on the estimation method (OLS and various instrumental variables). Dustmann and Fabri also report a positive effect on employment in the UK. The study on German fluency in Germany by Dustmann (1994) suggests a wage return of 7-15% using OLS with Heckman selection. On the other hand, Yao and van Ours (2015) find only modest wage effects for women and none for men in the Netherlands with respect to fluency in Dutch.

In the developing or emerging countries, a particularly well documented relationship is the economic benefit of knowing English in India. A substantial positive return has been reported at least by Azam et al (2013) and Chakraborty and Bakshi (2016). A somewhat different angle to the same question is provided by Shastry (2012) who shows that areas in India that have had a lower threshold for learning English, have grown faster due to opportunities provided by globalisation and information technology. In connection with this literature, we also show in this study that knowledge of Spanish language has a significant employment return among indigenous Mexicans (compared to those who do not know Spanish).

For this work, another relevant and interesting study is one by Munshi and Rosenzweig (2006) which shows that the choice of language in schooling in India has long-run implications for the labour market specialisation of the pupils as they grow up. In the study, they show that working class boys are disproportionately channelled to indigenous language schools which typically lead to traditional occupations, despite high returns to English language education. The rationale is that traditional occupations which depend on local job networks provide (or are perceived to provide) economic security. The results in our study can be interpreted in the same framework: Teaching the indigenous language to children may provide 'backup' job market opportunities in the traditional sector. This explanation is consistent with our finding that indigenous working age men who are bilingual are more likely to work than similar indigenous men who only know Spanish.

Another line of literature that this study contributes to, is the study of ethnic enclaves. Our results resonate for example with work by Edin, Frederiksson and Åslund (2003) and Damm (2009) who find that in Sweden, and Denmark, respectively, the labour market outcomes of ethnic minorities are

better if they live within their own enclaves. Neither Edin et al. or Damm explicitly study language, but it is likely that the use of a minority language among recent immigrants is one of the key factors in creating mutual understanding, job referral networks or other valuable information.

Compared to the existing literature, one of the main contribution of this study is to show that the logic of economic returns may also apply to minority languages. Studying this issue is very difficult for two reasons. Firstly, since many minority languages are in relative decline, the ex-ante view must be that these languages are associated with small, if any, economic benefits. Secondly, only few countries have a large enough number of minority languages for which the relevant data can be found. In our study, we are able to estimate the return to bilingualism for a total of 33 indigenous languages with a good sample size. This is a large enough group to not only measure the variability of economic benefits across the languages, but also to associate the economic return to the likelihood of intergenerational transmission of the language across the 33 groups.

The article is structured as follows. Section 2 gives an overview of indigenous languages in Mexico. In Section 3, we present an analysis of the employment return to bilingualism, for all indigenous languages together, and separately for each language. In Section 4 we present a basic model of language choice and transmission and continue with its empirical implementation in Section 5. Section 6 concludes. The data is introduced within sections 2, 3 and 5 as appropriate.

2. Indigenous Populations of Mexico

Mexico is integrated by a diverse mixture of native cultures. One manifestation of the richness of this diversity is the existence of 66 local languages which are spoken by 7.4 million people. Warnings about the risk of disappearance of these language enclaves are not new nor unwarranted in the context studied. It was estimated in 1889 that around 38% of the population spoke an indigenous language, down from 60% in 1820.[1] As of 2015, 21% of the population identified as indigenous yet only a third of them spoke a native language, that is 6.6% of the total population. Some fear that in the absence of institutional support, especially indigenous-language schooling, an important part of the cultural heritage of these millenarian cultures will be forever lost.

We observe that indigenous populations tend to be tightly concentrated in language enclaves and that this feature of their societal structuring is central for explaining the survival of their language over time. In fact, we find that while indigenous groups make up a relatively small share of the total population, in large parts of the country indigenous groups are highly concentrated and even surpass non-indigenous populations. The majority of the existing language enclaves are located at the centre and south of the country, and as shown below, we identify a total of 33 ‘language clusters’.

Figures 1 and 2 provide a current snapshot of the distribution of indigenous people and languages in Mexico, based on the Census data of 2015. Figure 1 shows the proportion of municipal population that identifies as indigenous. The second map in Figure 2 shows the proportion of people who speak an indigenous Mexican language. It is apparent that indigenous populations tend to be geographically clustered, and that the languages skills are at highest levels at the centres of these clusters.

[1] García Cubas (1904).

Figure 1. Proportion of the population that identifies as 'indigenous Mexican' by municipality, 2015.

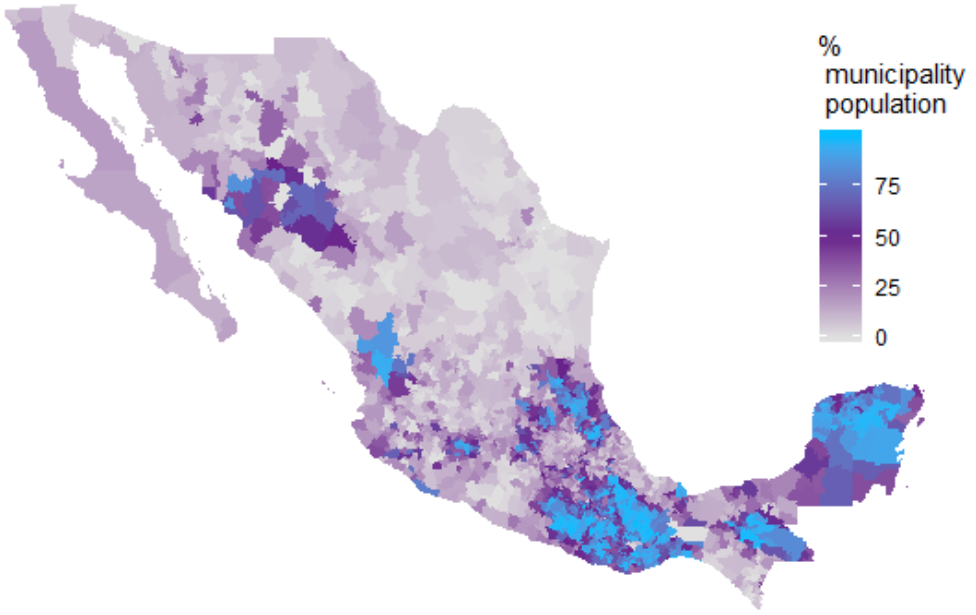
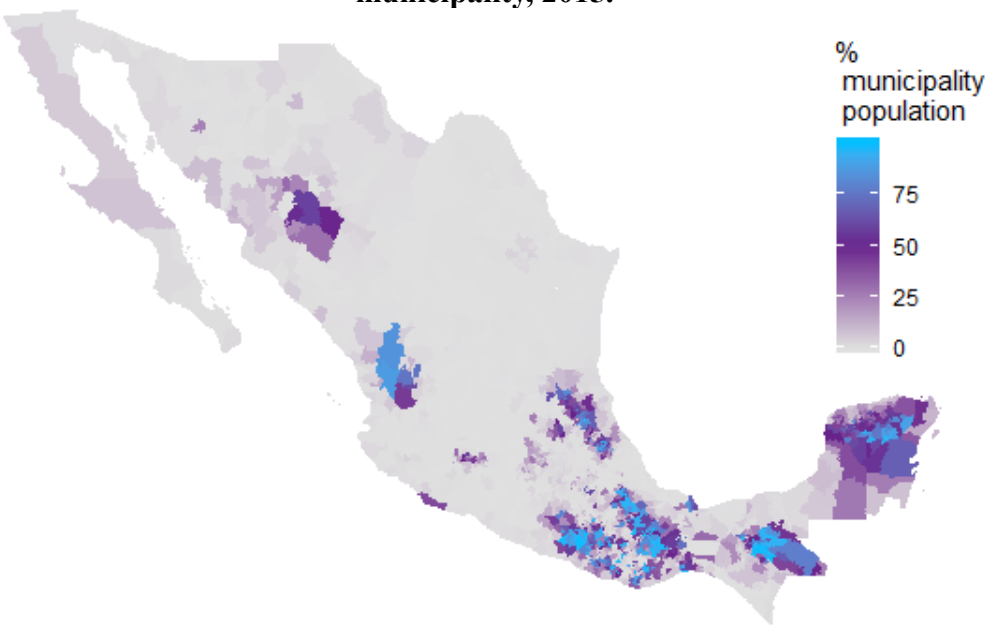


Figure 2. Proportion of the population that speaks an indigenous Mexican language by municipality, 2015.



Notes on Figures 1 and 2: Own calculations based on Mexican Census of 2015.

The indigenous languages can be classified into 11 independent language families.[2] Some languages are more dominant than others but they all exhibit high geographical concentration. Table 1 summarizes information about the identified languages enclaves. These are all clusters with more than 10 thousand people. In the table, the clusters have been ordered by number of speakers (column 1). The second and third columns give the total number and the share of population of each language group at a national level. The most spoken indigenous languages are Nahuatl and Maya, which account for 1.49% and 0.74% of the total population (column 3) but 23.2% and 11.6% of the indigenous speaking population.

[2] Instituto Nacional de Lenguas Indígenas.

Table 1: Indigenous Language Clusters

Language	Indigenous Speaking Population					# Municipalities	
	National		Cluster		Cluster		
	Total	%	Total	%	Intensity	Cluster	Language
Nahuatl	1559110	1.499	1194252	20.761	76.598	135	1467
Maya	775880	.746	742894	22.134	95.748	112	481
Tzeltal	516382	.496	473266	44.637	91.65	20	327
Tzotzil	454682	.437	395976	46.784	87.088	31	301
Mixteco	452368	.434	332084	45.951	73.41	116	991
Zapoteco	412928	.397	295708	53.45	71.612	137	968
Otomi	271524	.261	183308	13.12	67.51	24	593
Totocana	240624	.231	183186	43.842	76.129	32	567
Chol	233180	.224	192998	49.166	82.767	9	240
Mazateco	221038	.212	155368	65.706	70.29	28	530
Huasteco	159170	.153	123632	32.667	77.672	10	300
Mazahua	133176	.128	90010	17.102	67.587	7	350
Purepecha	128746	.123	100860	18.224	78.34	12	280
Tlapaneco	123056	.118	97486	48.19	79.22	9	231
Chinanteco	122558	.117	96946	32.907	79.102	20	390
Mixe	122284	.117	88640	72.587	72.486	23	491
Tarahumara	69828	.067	48230	31.608	69.069	9	174
Zoque	61454	.059	40790	36.343	66.374	10	174
Amuzgo	52944	.05	46808	41.969	88.41	5	128
Tojolabal	50220	.048	44552	34.149	88.713	2	64
Huichol	47926	.046	36268	8.224	75.674	6	170
Chatino	47442	.045	42250	61.984	89.056	9	146
Popoluca	45152	.043	39032	35.288	86.445	4	116
Tepehuano	42592	.04	33364	37.566	78.333	3	64
Mayo	35186	.033	27310	3.739	77.616	5	69
Cora	25688	.024	20338	58.889	79.173	1	46
Triqui	25404	.024	18538	3.811	72.972	5	132
Chontal de Tabasco	25086	.024	21508	2.618	85.737	3	16
Huave	18114	.017	15412	63.549	85.083	3	91
Yaqui	17938	.017	10666	7.737	59.46	1	58
Popoloca	16080	.015	10026	22.626	62.35	1	62
Cuicateco	11546	.011	7918	57.753	68.577	6	110
Pame	10788	.01	8834	25.182	81.887	2	36
Total	6603410	6.91	5218458	25.63	79.02		2456

The seemingly small share of language speakers as of the total population dissipates with a thinner geographical analysis. Of the total native speaking population, 79% live in areas in which: i. over 30% of the population speak the language; ii. over 10% of the national language population is concentrated in the municipality, or iii. more than 10,000 language speakers live in the municipality.[3] In these areas, 1 in 4 are native language speakers; at a national level they represent only 1 in 14.[4]

Language specific cases are often more dramatic, following the Maya example, while the speaking population represents 0.74% of the national population, it measures over a fifth (22.13%) of the cluster's population. Furthermore, 95.7% of all the Maya speaking population live in the cluster's

[3] 33 spoken languages with a total population of 73,316 have no identified cluster
[4] 25.63% -bottom of column 5.

municipalities (column 6). Column 6, the 'Cluster intensity', provides a measure of language dispersion as the share of language specific speakers living in what has been defined as the language cluster.

The seventh and eight columns give the number of language specific cluster municipalities and the total number of municipalities in which the language is spoken by anyone in the Census data. In the country there is a total of 2456 municipalities.

The map A1 in the Appendix aids the table by geographically locating cluster municipalities. Cluster-specific maps of the 16 languages with a population of over 100,000 speakers are in Figures A2-A4 in the Appendix. These maps effectively plot the municipality level estimates of an analogous variable to column 5 of Table 1. The maps show details of the proportion of the population within the municipality which speaks each indigenous language. The existence of these clusters is in the core of understanding how native languages have survived over time despite the lack of formal mechanisms for language preservation.

3. Labour Market benefits of bilingualism

In this section we estimate the employment effects of attaining indigenous bilingualism conditional on being of indigenous origin. The analysis is based on two different data sources; Mexican Censuses (from 2000, 2010 and 2015) and the National Household Income and Expenditure Survey (ENIGH) from 2016. The latter data is a more detailed employment and income survey at the cost of a smaller sample size.

We estimate the employment returns to languages in two steps. Firstly, we use all of the four datasets above, to estimate the wage and employment return to both indigenous languages and Spanish. The comparison groups are the monolinguals; in the first case the Spanish-only speaking indigenous people, and in the latter case the indigenous-only monolinguals. Secondly, we estimate the economic returns to each indigenous language separately. This can only be done in the Census, since the ENIGH data does not have large enough numbers of observations for each language group. Furthermore, estimation of wage returns using the Census are ridden with a large non-response to wages, and therefore we focus on the likelihood of being employed.

The Census contains information about ethnic group belonging and knowledge of ethnic language. Distinguishing between ethnic belonging and knowledge of language is only possible from the 2000 census onwards (2010, 2015) and is central to the matching estimator that we construct. All individual variables in our analysis (gender, birthplace, employment, schooling, household composition and age) as well as locality and municipality level characteristics, such as rural-urban status, comes from the census as well. Summary statistics for the census samples are presented in Appendix Table 7, and the corresponding summary statistics for the ENIGH data are in Appendix Table 8.

We restrict the sample to non-migrant indigenous males between the age of 25 and 64. One desirable feature of the Census data is that we observe state of birth and residence 5 years prior to the census interview so that our estimates look at the effects of bilingualism of long-term residents. Measuring migration across municipalities is not possible since the census questionnaire only asks about state of birthplace prior state of residence. To some extent, this sorts out the problem that arises from the fact that localities and cities often span several municipalities. After the inclusion of all the restrictions on the data, we are left with a sample of 358,347 individuals for our estimates of 2015, 419,964 for 2010 and 136,416 for 2000. Using similar restriction in the ENIGH data, we are left with 4,521 observations.

The strategy to estimate the effect of bilingualism uses a combination of matching and OLS estimation. In the first step, we restrict the sample tightly to families in which there is at least one indigenous speaker, in a way that the indigenous language corresponds to the main indigenous language within the municipality. This restricts the sample to indigenous people who tend to live within their own language clusters.

Within that sample we select males aged 25 to 64, who report being indigenous and who are non-migrants. This group is then divided into treatment depending on whether they are monolingual Spanish speakers (control), or bilingual (treatment). As even the monolinguals in the sample live in households with at least one indigenous speaker, this matching guarantees that the control group has only a minimal social distance to the treatment group.

In summary**:**

$$d_i = 1 \begin{cases} \text{indigenous speaking} = 1 \\ \text{indigenous belonging} = 1 \\ \text{indigenous speaking family} = 1 \\ \text{family language} = \text{mun language} \\ \text{age= (25 - 64)} \\ \text{non-migrant=1} \\ \text{sex=male} \end{cases} , d_i = 0 \begin{cases} \text{indigenous speaking} = 0 \\ \text{indigenous belonging} = 1 \\ \text{indigenous speaking family} = 1 \\ \text{family language} = \text{mun language} \\ \text{age= (25 - 64)} \\ \text{non-migrant=1} \\ \text{sex=male} \end{cases}$$

In the second step, we recover the likelihood of employment with a linear probability model summarized as:

$$employment_i = \gamma_0 + \lambda * d_i + \Lambda X_i + \varepsilon_i$$

where d_i is the matching identifier between the treatment group (bilingual) and the non-treated (non bilingual); X_i corresponds to a fourth-order polynomial of age, schooling and locality size controls in addition to municipality level fixed effects. The inclusion of municipality fixed effects together with the matching in the first step guarantees that the employment comparison between mono- and bilinguals is done within the same geographic area, within people who identify with the same indigenous group and language. The parameter λ is the estimate of interest and it corresponds to the increased likelihood of employment for bilingual working age men.

As we present the results, we show results with and without the 1st step matching, to show its effect on the estimates. A priori, there's a reasonable expectation that without the 1st step matching, the control group of monolinguals would include Spanish speakers who have larger social distance to the treatment group, who are also is likely to have higher socioeconomic status. If that is the case, it would bias the employment effect of bilingualism downwards.

We also report the employment return to knowing Spanish. To do that, the control group is changed to indigenous males who only speak the indigenous language but not Spanish.

Results using Census data for 2000, 2010 and 2015 are summarized in Table 2. These are the set of estimates with full set of controls which include age (linear and non-linear effects), locality size controls, schooling years and municipality fixed effects. The first row of results corresponds estimates of returns to employment without the 1st step matching, so that we include all indigenous working age males, regardless of whether the individual lives in an indigenous speaking household, what is

the language spoken in the municipality and whether the individual has migrated. Employment returns for this naive estimation are still positive for the two first periods studied and zero for 2015.

Table 2: Returns of Bilingualism: Increase in likelihood of Employment (Indigenous Populations: 2000, 2010, 2015)

	λ Employment Likelihood		
	2000	2010	2015
Indigenous Bilingual: Fixed Effects			
Learn Indig	0.017***	0.015***	0.000
	(0.002)	(0.001)	(0.001)
N	157991	634627	762538
adj. R^2	0.120	0.085	0.112
F	157.5	763.2	2145.4
Indigenous Bilingual: Matching			
Learn Indig	0.035***	0.030***	0.019***
	(0.004)	(0.002)	(0.002)
N	136416	419964	358347
adj. R^2	0.122	0.099	0.134
F	121.8	319.9	508.3
Spanish Bilingual: Matching			
Learn Spanish	0.015***	0.014***	0.037***
	(0.003)	(0.0018)	(0.002)
N	147439	449784	350617
adj. R^2	0.125	0.105	0.143
F	136.9	317.4	499.4

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

The second row of Table 2 corresponds to the 2-step estimates which include matching. Now, the estimates capture the difference in the probability of employment between indigenous bilingual and indigenous Spanish monolingual individuals - and other constraints described in λ. The critical assumption is that the employment probability differentials between the compared groups is driven by bilingualism, and not unobservable factors.

The last row in these tables summarizes the work returns from learning Spanish. These are estimates from a similar matching grouping as in the second row. Now the compared groups differ in that monolingual individuals are those that only speak indigenous; hence, these estimates correspond to the economic benefits of learning Spanish. Overall, in each Census year, we find that bilinguals

outperform both Spanish and indigenous language monolinguals, given the educational and other characteristics. There is a slight downward trend in returns on knowing an indigenous language and an upward trend in returns for knowing Spanish.

Appendix Table 10 reports the equivalent estimates using the 2016 National Household Income and Expenditure Survey (ENIGH). Using a combination of matching and OLS, the employment return to bilingualism is found to be a statistically significant 2.7 percentage points, which is not wildly different from the 1.9 percentage points found in the 2015 Census. Using ENIGH, we were also able to estimate the wage return to bilingualism. We found a positive, but not a statistically significant effect for indigenous languages.

Table 3 recovers employment returns for the ten largest indigenous languages. These estimates correspond to the baseline matching estimator with full set of controls for each of these languages. From it, a high degree of heterogeneity between language groups becomes apparent, with some language returns exceeding in threefold the overall indigenous effect of Table 2. Interestingly, one language exhibits negative returns to bilingualism: Mixteco. Further analysis of this language shows a geographic overlap between the Mixteco and Zapoteco and other language enclaves. Historical records document rivalry between the two main groups from pre-columbine periods, this raises the question of whether competition between groups could be behind this outcome.

Table 3: Language Specific Returns to Bilingualism: Increase in likelihood of Employment: 2015)

Employment Likelihood	
Language	λ
Chol	.106***
	(.019)
Tzotzil	.070***
	(.017)
Tzeltal	.065***
	(.020)
Mazateco	.044**
	(.020)
Maya	.038***
	(.004)
Mazahua	.036**
	(.016)
Zapoteco	.031***
	(.008)
Nahuatl	.022***
	(.005)
Totocana	.002
	(.013)
Otomi	−.012
	(.011)
Mixteco	−.018*
	(.018)

Standard Error in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Overall, we estimate the employment return to 33 indigenous languages. This is the maximal number of languages we can include if we require that both the treatment and the control group must have at least 60 individuals in the 2nd step estimation. Table 9 in the Appendix provides language specific returns and summary statistics for all languages.

4. Transmission of language

When an individual remains monolingual, direct communication links are limited to other monolingual speakers and bilingual speakers that know his or her own language. In a rational world a native speaker of a particular language will chose to learn another language if the utility gain derived from increasing communication links outweigh the costs of learning (a simple model of bilingualism, but not fully applicable to our study, is provided by Church and King, 1993).

In bilingual or multilingual environments and families, roughly the same idea applies to the efforts of parents to teach their children a particular language. Parents may master a menu of languages, and associate languages with different expected long term social and economic benefits. The costs of teaching a particular language to children may also vary greatly depending on the availability of speakers in the household and the exposure of the children to the language in the local environment.

Suppose that the utility derived from knowing a language is increasing in the number of speakers and that the costs of learning the language decrease with higher exposure to the language. Were this the case, it follows that the efforts to learn a language is an increasing function of the number of potential speakers. Language networks have cost reducing externalities too. This is because the costs of acquiring the language are a decreasing function of language exposure which is largely determined by the network characteristics.

An important feature of Mexico and its indigenous languages is that bilingual school education is very underdeveloped. Despite legislative efforts to increase indigenous language education in primary schooling, the introduction of bilingual education has been variable. As such, Spanish remains the *de facto* language of education, and exclusively so at the secondary and higher levels of education. For further details on the institutional environment, please refer to Appendix 2.

We model the social and economics networks that determine the transmission of languages as follows:

1. Family network:

The costs of teaching a language to a child declines with more adults in the household who are able to speak it. Furthermore, the existence of relatives who speak an indigenous language also increase the social benefits of knowing the language. In the estimates below, we show that the transmission is higher when both parents can speak the minority language, than in mixed couples where only one of the parents can speak it. Additional extended family members in the household who speak the minority language also increase the likelihood of transmission.

2. Local Area Network:

The larger the proportion of minority language speakers there is in a municipality, the larger the likelihood that the language is taught to children within households. This is one of the most important determinants of bilingualism in Mexico, and makes sense both in terms of the costs of learning a language and the benefits of knowing it.

3. Economic return to bilingualism:

Commanding an additional language is a skill, with a potential positive economic return. In the context of Mexico, the economic benefits from native languages are likely to exist due to employment networks in certain professions, especially agriculture, which is a key form of livelihood among indigenous populations of Mexico (for evidence, please refer to the summary statistics accompanying the Appendix Table 9)

Our main hypothesis is that the decision to pass the language is partly informed by the perceived economic opportunities that the language skill may provide to their children. The perceived economic return are the existing economic returns that the parents' generation has enjoyed.

Overall, we build a simple model of intergenerational language transmission as a function of the social networks and the economic benefits. The key elements of the family-level model are as follows:

P(Language Passed|X) = F_1(Language resources in the family) +
F_2(Language resources in the local community) +
Economic return to the language +
Family and local controls +
Language-level controls +
Error term

The study focuses on families with both parents present. In such cases, a language resource in the family (Function F_1) consist of whether both or only one of the parents can speak the native language. We also take into account whether there are other adults in the household who can speak the language (such as grandparents). We assume (and test) that each additional adult who can speak a native language in the household, generally increases the likelihood that the language is passed down to the next generation.

For function F_2, which measures the strength of the language in the local area, we compute the proportion of people in the municipality who speak the same minority language as the household does. In practice, this measures the potential interactions that can be made using the minority language in the local area. In our data of bilingual families, the average family is located in a municipality where 52% of the local population can speak the same language as family. This suggest that the typical bilingual family lives in a 'core' of the minority language. On the other hand, this measure has a large variation, showing that the strength of the local language network cannot in general be taken for granted, and it is important to control for it.

4.1 Data and sample

The data is based on the Mexican Census of 2015. The sample is limited to the household respondent and his/her spouse and children. Families which speak only Spanish are excluded, so that at least one of the parents states that they can speak a native Mexican language. To simplify analysis, single-parent families and families where parents speak two different native languages are excluded. As such, each bilingual nuclear family is categorised to belonging to one of the native Mexican language groups. Further, the age of the mother has been restricted to range 25-54.

Table 4 presents the summary statistics on the 2015 sample of households. Within the sample, 64.5 percent of parents have passed the minority language to their children.[5] In 9 percent of the families, only mother can speak the native language, and in 12 percent, only father. This implies that in 79 percent of the families, both parents state that they can speak a native language. We have not

5 In 93% of the families, either all or none of the children learn the minority language. Therefore we have rounded the share of children who speak the language to either 0 or 1. The language skills of children under 4 years are not defined.

documented the Spanish skills, since it is increasingly rare that people in Mexico can't speak any Spanish. All children are exposed to Spanish by the school system.

A noteworthy fact in Table 4 is that only about two-thirds of children indigenous-speaking parents learn the indigenous language. This goes to show that a large fraction of families is likely to be 'in the margin' of deciding whether to pass the indigenous language to the next generation.

Table 4 also lists a number of key household variables that may affect the transmission of language within the household. The table also includes variables at the native language group level and the municipality level. At language group level, the main variable of interest is the group-specific employment return to bilingualism, or the estimated increase in likelihood of employment from being able to speak the native language in addition to Spanish. Other variables that proxy the economic importance of the group are the group size, as well as the average wealth and education in the group. All of these variables have substantial variation across the 33 groups covered by the sample.

Table 4. Summary statistics

Variable	Obs	Mean	SD	Min	Max
Household variables:					
Children speak native	225,183	0.646	0.478	0	1
Only mother speaks native	225,183	0.093	0.290	0	1
Only father speaks native	225,183	0.120	0.325	0	1
# Other HH adults speak native	225,183	0.145	0.445	0	11
% municipality share HH lang.	225,183	0.521	0.314	1.5E-06	9.7E-01
Mother's years of educ.	225,183	4.822	3.879	0	18
Father's years of educ.	225,183	5.468	3.917	0	18
Mother's age	225,183	38.751	8.009	25	54
Father's age	225,183	42.527	9.751	12	100
Normalised HH wealth	225,183	0.042	1.009	-1.481	3.483
Urban household	225,183	0.311	0.463	0	1
Language group variables (n=33):					
Group rate of return to bilingualism	225,183	0.028	0.033	-0.048	0.104
Group size (# of households)	225,183	236002	209959	2362	596636
Average wealth in group	225,183	0.314	0.439	-0.538	1.151
Average yrs. of education in group	225,183	4.960	0.659	3.131	7.261
Municipality variables (n=1962):					
Educational deprivation index	225,183	33.062	10.633	5.1	60.6
Health deprivation index	225,183	13.752	6.722	0.9	77.4
Housing deprivation index	225,183	32.188	16.935	1.3	82.7
Food deprivation index	225,183	27.960	12.311	0.5	85.7

4.2 Results

Table 5 shows results on a number of estimations for the determination of language transmission, using family and municipality characteristics, and the economic return to languages.

The first column is the benchmark model for language transmission, and it uses only the household characteristics, as well as regional fixed effects. The first important result is that if either mother or father can't speak the native language, it is much more likely that the language is not transmitted to children. If father doesn't speak the language, the likelihood of transmission falls by 44 percentage

points. Mother's ability to speak the minority language is estimated to be somewhat more important than father's (47 percentage points), which is consistent with mothers spending more time with their children than the father.

Additional adults in the household that can speak the minority language increase the likelihood of language transmission by about 4.5 percentage points per person. While the effect is statistically quite significant, the size of the effect on children is only about 1/10 of the effect of a parent's language.

Table 5. Return to native language and its transmission

Dependent: Children can speak native language				
	[1]	[2]	[3]	[4]
Return to bilingualism		.315+	.315+	.461*
		[.174]	[.174]	[.18]
Urban x Return to bilingualism				-0.452
				[.287]
Urban				-.0383*
				[.0143]
Only mother speaks native	-.441**	-.441**	-.44**	-.438**
	[.0327]	[.0326]	[.0324]	[.0332]
Only father speaks native	-.472**	-.472**	-.471**	-.468**
	[.0341]	[.034]	[.034]	[.0349]
# Other adults speak native	.0452**	.0456**	.0455**	.046**
	[.00543]	[.00531]	[.00531]	[.00545]
% municipality share HH lang.	.438**	.437**	.445**	.444**
	[.0436]	[.0435]	[.0426]	[.0429]
Mother's years of educ.	-.00725**	-.00719**	-.00739**	-.00765**
	[.000692]	[.000692]	[.000725]	[.000711]
Father's years of educ.	-.00411**	-.00412**	-.00438**	-.00433**
	[.000585]	[.000584]	[.00063]	[.000604]
Mother's age	0.000405	0.000421	0.000388	0.000403
	[.000322]	[.000324]	[.000327]	[.000352]
Father's age	-.000602**	-.000596**	-.000632**	-.000709**
	[.000206]	[.000203]	[.000197]	[.000185]
Normalised HH wealth	-.0516**	-.0519**	-.0515**	-.0429**
	[.0049]	[.00469]	[.00441]	[.00332]
Municipal controls:				
Educational deprivation			-.00142*	-.00163**
			[.000576]	[.000589]
Health deprivation			-0.0000788	0.000205
			[.000657]	[.000574]
Housing deprivation			0.000382	0.000391
			[.00033]	[.000323]
Food deprivation			.000578**	.000649**
			[.000197]	[.000217]
Constant	.579**	.569**	.589**	.605**
	[.0418]	[.0427]	[.0417]	[.0424]
Region fixed effects	Yes	Yes	Yes	Yes
Observations	225,183	225,183	225,183	225,183
R-squared	0.523	0.523	0.524	0.526

Notes: Linear probability. If at least 50% of children speak native, the family is coded as 1 in the dependent variable. Standard errors clustered at language group level, ** p<0.01, * p<0.05, + p<0.1.

With regards to the local network of minority language speakers, the first column of Table 5 suggests that if the local proportion of minority language speakers increases by 10 percentage points, the likelihood of transmitting the language in the household increases by 4.4 percentage points, which is not far from the effect that one additional adult speaker in the household has. This is a variable that has substantial variability across households, with a standard deviation of 0.3. This implies that moving a bilingual family to a municipality with 1 SD larger share of minority speakers would imply a 13.2 percentage points (0.3*0.44) larger likelihood that the language is passed to the next generation.

Parental education, age and household wealth (based on an index of items) all have a negative and significant association on the likelihood of language transmission. Of these, it is worth noting that each year of maternal education reduces the likelihood of the language transmission by about 0.7 percentage points, and one standard deviation of household wealth by about 5 percentage points. An explanation for these effects could be that further study, typically conducted in Spanish, gears the parents to overlook the potential value of the minority languages. The effect of wealth and also education suggest that in general the Indian languages are strongly associated with lower socio-economic status in Mexico, a fact that will have to be kept in mind when interpreting the results and their causality.

Columns 2 adds the employment return of the language-specific bilingualism into the model. The effect of the economic return in itself suggests a positive, marginally statistically significant effect (at 10% level). Here it is important to note that since this variable varies by the 33 native groups, the standard errors are clustered by these groups. Column 3 further adds controls for the municipal level of economic deprivation. Since there are nearly 2000 municipalities covered by the sample, adding these controls allows us to control for sources of potential omitted variable bias in the model. The variables included are indices for educational, health, housing and food deprivation. Remarkably, these have very little effect on the results of interest, suggesting that local levels of economic development do not bias the result. [6]

The fact that a very large proportion of the indigenous males work in agriculture (see Appendix Table 9), suggests that the employment return to the native languages must be largely driven by employment networks in agriculture. If that is the case, it is possible that the families in rural areas respond to this economic benefit more than in urban areas, by making sure their children learn the indigenous language. This is why, in the final column of Table 5, we have interacted the employment return with urban location. The results show that in rural areas, the higher employment return is associated with increased likelihood to pass the language at 5% statistical significance level, whereas in the urban areas, the effect is very close to zero (.461 - .452 = .009). In rural areas, the size of the effect is not trivial: If the employment return to bilingualism increases by 2 standard deviations (2*.033), the likelihood of passing the language increases by 3 percentage points (2*.033*.461 = .0304).

4.3 Robustness checks

6 An alternative to municipal multidimensional deprivation would be to use municipal fixed effects. The problem with this approach is that since the native groups are highly regional (see the appendix maps), municipal fixed effects would not have sufficient variation in most of the country, but would instead be based on the largest cities which host multiple indigenous groups, but with few, and very selected individuals. This would not give the representative estimates we are looking for.

An obvious concern that arises from the estimates of Table 5 is that the result on the employment return observed in Table 5 is in fact reflecting the generic socioeconomic status of the language. It may be that estimates of the return to bilingualism are not actually returns on skills per se, but that they are signals of the relative prestige of the language: Workers who belong to a higher status group, are more likely to find work (leading to the variability in employment return across groups). They are also more likely to pass the language to their offspring as a 'signal' of the group membership.

Due to this concern, in Table 6 we estimate how the inclusion of other variables that measure the status of the group change the estimates. If the generic group status is a source of omitted variable bias, inclusion of these variables should reduce the estimate on how much the return to bilingualism affects language transmission.

The columns 1-3 of Table 6 include, consecutively, (1) the size of the group as measured by the logarithm of the number of households in Mexico where the language is spoken, (2) The average wealth index of the households of the group and (3) the average education of the households in this group. Further, in column 4, all of these variables are included at the same time.

Remarkably, the results show that the effect of employment return on the language transmission appears to be fairly orthogonal to these variables. If anything, the result strengthens by a small margin. From this, we can conclude that the relative socioeconomic status of the languages is not driving the main result in Table 5.

Table 6. Robustness check: Language group status

Dependent: Children can speak native language				
	[1]	[2]	[3]	[4]
Return to bilingualism	.284+	.325+	.339*	.295*
	[.146]	[.165]	[.153]	[.143]
Ln Group size	-.0108*			-.00919*
	[.00412]			[.00395]
Group avg. wealth		-.0307+		-0.0116
		[.0154]		[.0208]
Group avg. years of education			-0.013	-0.00117
			[.00896]	[.0117]
Household controls	Yes	Yes	Yes	Yes
Municipal controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Observations	225,183	225,183	225,183	225,183
R-squared	0.524	0.524	0.524	0.524

Notes: All models include the same controls as in column 4 of Table 2. Linear probability. Standard errors clustered at language group level, ** p<0.01, * p<0.05, + p<0.1.

5. Conclusions

This study explores the labour market aspects of minority languages in the context of Mexico, a country with a rich heritage of indigenous languages. The paper provides two linked results. The analysis is based on Mexican Censuses and the National Household Income and Expenditure Survey of 2016.

Firstly, we estimate the economic return to being bilingual for a large number of indigenous Mexican languages. We show that on average, similar indigenous Mexican males are more likely to be

employed if they can speak both the indigenous language and Spanish as opposed to Spanish only. The employment benefit varies by language and can be explained by the domination of agriculture by the indigenous groups. The result is based on a combination of matching and OLS estimators.

Secondly, we show that the larger the employment return there is to the skill in a particular language, the more likely it is that parents transmit the language skills to their children. In is analysis, we build a detailed picture of bilingual households in Mexico, and model all key factors in the family and the local environment that affect the transmission of indigenous languages either by reducing the cost, or increasing the benefits of knowing them.

The results provide a new opening for the existing literature that estimates returns to language skills by showing that the returns can also be estimated for minority languages that lack economic clout. The study also provides a unique systematic documentation on how economic factors can affect the continuation and survival of minority languages that lack the support of official institutions in developing countries.

The results also provide links to the literature on the labour markets of ethnic enclaves. It is also apparent from the results that additional language skills can be thought of as forms of insurance, that allow the speakers to access additional niche labour markets.

References:

Azam, Mehtabul, Aimee Chin, and Nishith Prakash (2013). The Returns to English-Language Skills in India. *Economic Development and Cultural Change*, 61(2): 335-367.

Bleakley, Hoyt and Aimee Chin (2004). Language Skills and Earnings: Evidence from Childhood immigrants. *The Review of Economics and Statistics*, 86(2): 481-496.

Chakraborty, Tanika, Shilpi Kapur Bakshi (2016). English language premium: Evidence from a policy experiment in India. *Economics of Education Review*, 50, Feb 2016, pp. 1-16.

Chiswick, Barry R. and Paul W. Miller (1995). The Endogeneity Between Language and Earnings: International Analyses. *Journal of Labor Economics*, 13(2): 246-288.

Church, Jeffrey and Ian King (1993). Bilingualism and Network Externalities. *The Canadian Journal f Economics*, 26(2): 337-345.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2): 281-314.

Dustmann, Christian (1994). Speaking fluency, writing fluency and earnings of migrants. *Journal of Population Economics*, 7(2): 133-156.

Dustmann, Christian and Francesca Fabbri (2003). Language Proficiency and Labour Market Performance of Immigrants in the UK. *The Economic Journal*, 113(489):695-717

Edin, Per-Anders, Peter Fredriksson and Olof Åslund (2003). Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment. *The Quarterly Journal of Economics*, 118(1): 329–357.

García Cubas, A. (1904). El libro de mis recuerdos. *Imprenta de Arturo García Cubas, Hermanos Sucesores, México*.

García Segura, Sonia (2004). De la Educación Indígena a la Educación Bilingüe Intercultural (La comunidad p'urhépecha, Michoacán, México). Investigación temática - educación intercultural. 9(22), p. 61-81.

Instituto Nacional de las Lenguas Indígenas (2008). Catálogo de las Lenguas Indígenas Nacionales. Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas.

Jiménez-Naranjo, Y. and Mendoza-Zuany, R. (2016). La educación indígena en México: una evaluación de política pública integral, cualitativa y participativa. LiminaR, 14(1), 60-72.

Miranda, Alfonso and Yu Zhu (2013). English deficiency and the native-immigrant wage gap. *Economics Letters*, 118(1): 38-41.

Munshi, Kaivan and Rosenzweig, Mark (2006). Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy. *American Economic Review*, 96(4): 1225-1252.

Nettle, Daniel, and Suzanne Romaine (2002). *Vanishing Voices: The Extinction of the World's Languages*. Oxford University Press, USA.

Recondo, David (2007). La política del gatopardo. Multiculturalismo y democracia en Oaxaca. Capítulo 2. El surgimiento de un actor político indígena México, Publicaciones de la Casa Chata. p. 83-120. https://books.openedition.org/cemca/2066?lang=es

Salmerón Castro, F. and Porras Delgado, R. (2010). Los Grandes Problemas de México. Part 4, Chapter 17,"La educación indígena: fundamentos teóricos y propuestas de política pública". El Colegio de México, Vol VII, p. 509-546.

Sarmiento Silva, Sergio (1985). El Consejo Nacional de Pueblos Indígenas y la política indigenista. Revista Mexicana de Sociología. Universidad Nacional Autónoma de México 47(3) p.197-215.

Schmelkes, Sylvia, (2006). La interculturalidad en la educación básica. Proyecto Regional de Educación para América Latina y el Caribe, UNESCO. Num. 3. p. 120-127.

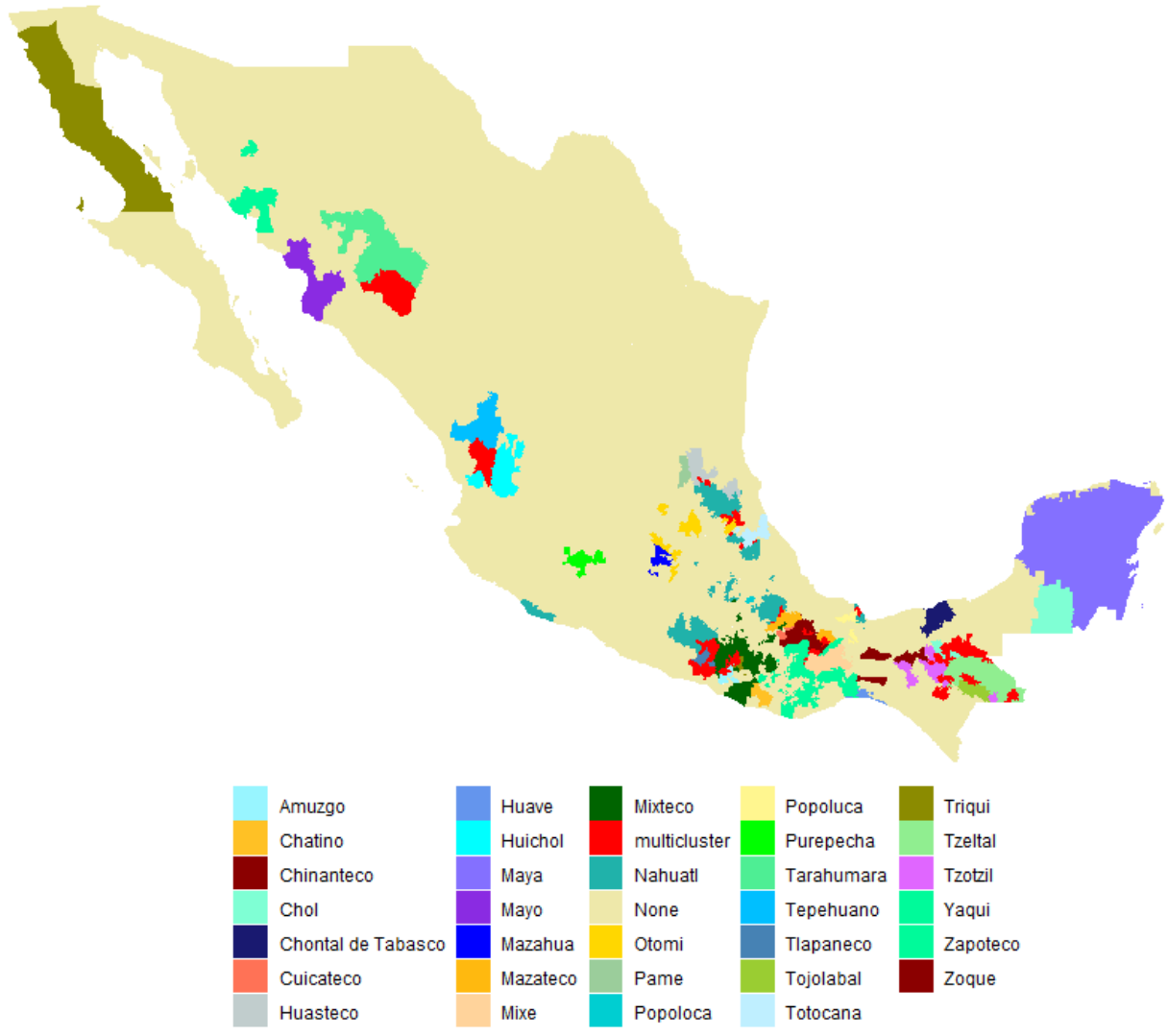
Shatry, Gauri Kartini (2012). Human Capital Response to Globalization. Education and Information Technology in India. *The Journal of Human Resources*, 47(2):287-330

Stavenhaven, Rodolfo (1988). Anuario de etnología y antropología social. Chapter 7: Los derechos humanos de los pueblos indios. Vol 1, México, D.F., p. 130 - 136.

Yao, Yuksin and Jan C. van Ours (2015). Language skills and the labor market performance of immigrants in the Netherlands. *Labour Economics*, 34(June 2015): 76-85.

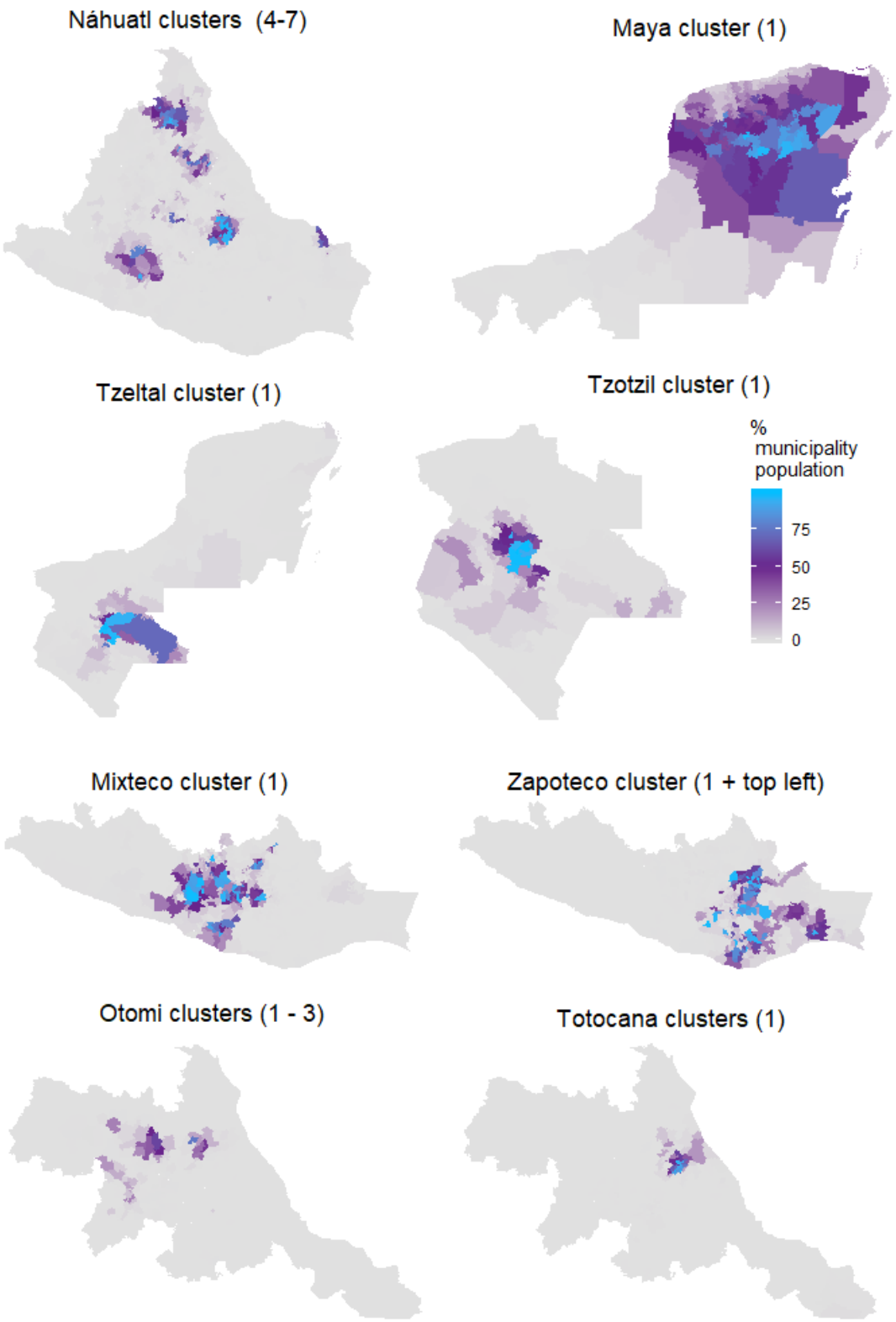
Appendix 1: Other Figures and Tables

Figure A1 The geographic location of specific indigenous language clusters in Mexico, 2015



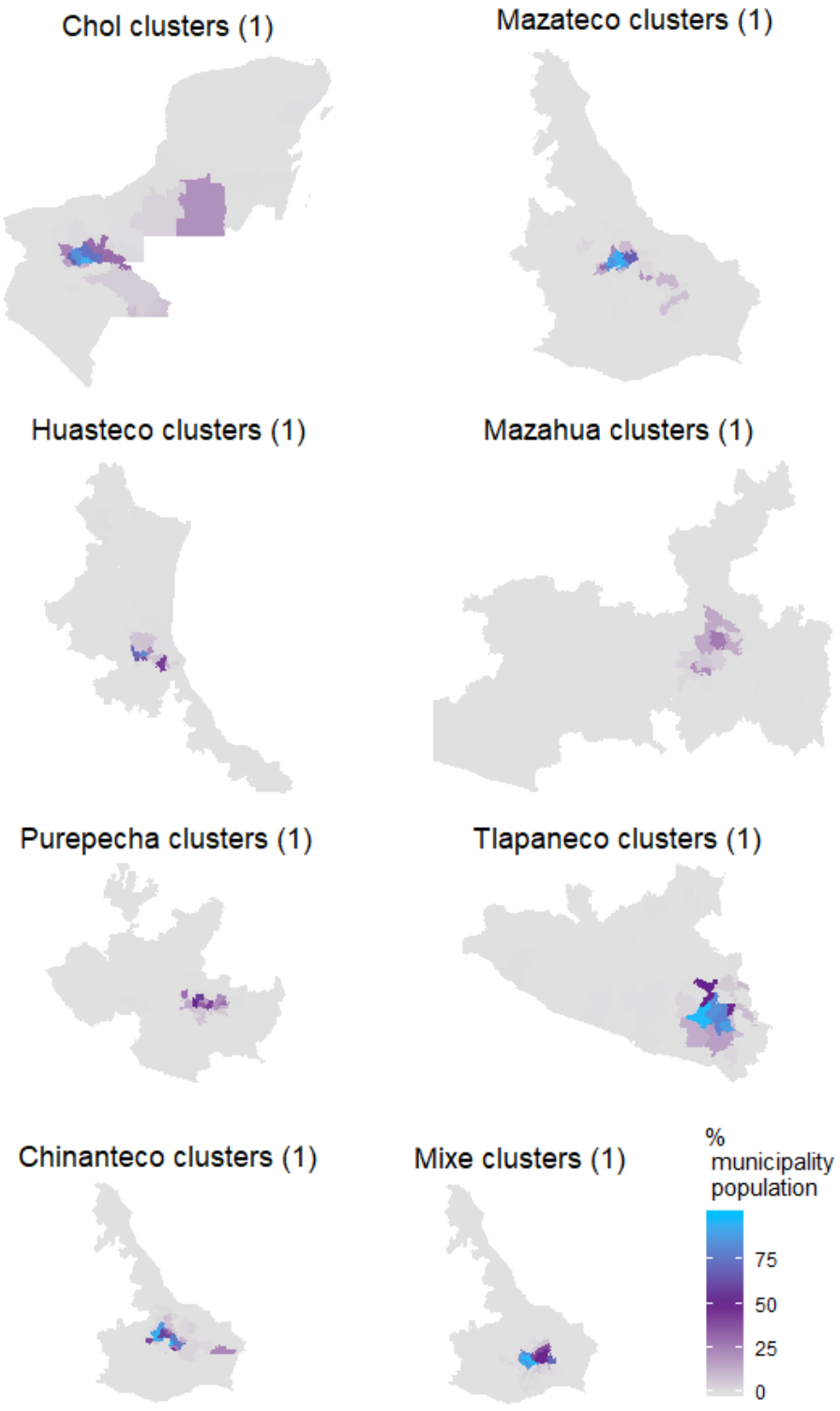
Notes: Own calculations based on Mexican Census of 2015.

Figure A2 Detailed maps for 16 largest language clusters, part 1/2.



Notes: Own calculations based on Mexican Census 2015.

Figure A3 Detailed maps for 16 largest language clusters, part 2/2.



Notes: Own calculations based on Mexican Census 2015.

Table 7. Summary Statistics Census 2015 by population groups

Indigenous	Speaks Indigenous	Average Age	Illiteracy Rate	School Years	Work Status	Sample Size
All non-migrant male, 25-64						
21.3	6.8	41.4	0.04	9.60	0.85	2,352,008
All non-migrant male, 25-64 Identified with ethnic/indigenous group.						
100	28.3	41.6	0.07	8.03	0.83	799,977
All non-migrant male, 25-64 Identified with ethnic/indigenous group. Speaks Indigenous Language.						
100	100	42.3	0.16	6.02	0.80	357,369

* Estimates for three different samples. Non explicit variables as follow:
Indigenous: Share of population that self-identifies as indigenous
Schooling: Average years of schooling
Illiteracy Rate: Fraction that " does not know how to read or write a message"
Work Status: Fraction of the population working

Table 8. Summary Statistics ENIGH 2016, by population groups

Indigenous	Speaks Indigenous	Average Age	Illiteracy Rate	Post Primary	Work	Sample Size
All non-migrant male, 25-64						
0.30	0.06	41.9	0.04	0.29	0.91	56,217
All non-migrant male, 25-64 Identified with ethnic/indigenous group.						
1	0.20	42.1	0.07	0.39	0.92	17,776
All non-migrant male, 25-64 Identified with ethnic/indigenous group. Speaks Indigenous Language.						
1	1	42.9	0.15	0.60	0.94	3,968

* Estimates for three different samples. Non explicit variables as follow:
Indigenous: Share of population that self-identifies as indigenous
Post Primary: Share of population with less than primary school
Illiteracy Rate: Fraction that " does not know how to read or write a message"
Working: Fraction of the population working

Table 9. Employment Returns to Bilingualism by Language

Population: 25-64 Year Old Non-Migrant Males.

Language	λ	SE	Obs	Average School	Average Age	Agro Share
Chol***	0.106	0.019	13716	6.3	40.6	0.73
Mixe***	0.103	0.034	8808	6.2	41.5	0.47
Popoloca*	0.095	0.053	675	4.6	40.5	0.35
Mame	0.095	0.072	224	4.6	43.3	0.60
Pame	0.089	0.061	703	4.2	41.5	0.55
Cora	0.078	0.055	2008	4.9	40.6	0.48
Popoluca*	0.072	0.040	1915	5.1	41.7	0.67
Tzotzil***	0.070	0.017	28677	4.9	39.4	0.62
Chontal de Tabasco*	0.066	0.036	644	8.6	43.6	0.27
Zoque***	0.066	0.024	3956	5.4	41.9	0.65
Tzeltal***	0.065	0.020	25589	5.5	39.7	0.73
Huave	0.062	0.043	1902	6.6	42.2	0.54
Tojolabal	0.053	0.041	1596	4.8	40.3	0.82
Amuzgo	0.048	0.035	3948	4.5	40.5	0.58
Mazateco**	0.044	0.020	15254	5.4	41.6	0.48
Maya***	0.038	0.004	46780	7.1	43.3	0.24
Mazahua**	0.036	0.016	3521	5.8	44.7	0.22
Zapoteco**	0.031	0.008	32077	7.3	43.7	0.33
Nahuatl***	0.022	0.005	77750	6.3	42.5	0.34
Chontal de Oaxaca	0.021	0.047	473	6.3	47.0	0.66
Purepecha	0.018	0.014	4631	6.6	41.8	0.29
Huasteco	0.014	0.020	6386	6.5	42.6	0.24
Chinanteco	0.007	0.017	9432	6.1	42.5	0.64
Mayo	0.007	0.022	1083	7.6	46.3	0.37
Totocana	0.002	0.013	14330	5.9	43.5	0.49
Yaqui	-0.012	0.044	470	8.3	42.9	0.26
Otomi	-0.012	0.011	11671	6.5	44.2	0.25
Cuicateco	-0.015	0.033	1346	5.2	43.9	0.70
Tlapaneco	-0.017	0.034	8145	6.3	40.1	0.56
Mixteco*	-0.018	0.011	30013	5.8	42.6	0.41
Chatino*	-0.038	0.023	3815	4.3	41.8	0.60
Tepehuano	-0.045	0.034	3137	5.3	39.3	0.42
Tarahumara*	-0.048	0.026	5217	4.9	42.1	0.36
+Pop 25-64				9.8	41.3	0.10

Notes: *** P<0.001, ** P<0.05, * P<0.1

+ Pop 25-64: corresponds to non-migrant males (indigenous and non-indigenous).

Table 10: Returns on Employment and Income - ENIGH 2016

	μ Earnings Returns	λ Employment Returns
Indigenous Bilingual: Fixed Effects		
Learning Ind	-0.0893***	0.0269***
	(0.0255)	(0.00670)
N	13674	17651
adj. R^2	0.292	0.050
F	94.24	39.05
Indigenous Bilingual: Matching		
Learn Indig	0.0662	0.0273**
	(0.0481)	(0.0110)
N	3299	4591
adj. R^2	0.300	0.033
F	19.20	6.385
Spanish Bilingual: Matching		
Learn Spanish	0.431**	0.0564**
	(0.151)	(0.0234)
N	2740	3964
adj. R^2	0.316	0.049
F	17.00	6.031

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Appendix 2: Institutional and Education Policies for Indigenous Populations in Mexico

Mexican society is a collection of groups with profoundly diverse backgrounds. One manifestation of this is the mosaic of languages that exist. While a significant number of languages have been lost through policies of cultural homogenization that began during colonial times and persisted into well the second half of the twentieth century, as of 2015, there were 66 Mexican languages written and spoken in different parts of the country (7.4 million people).

Indigenous language speaking populations tend to be concentrated in tight geographical areas, a pattern that appears to be fundamental for the transmission of language, one of the core points studied in this research.

From the end of the 19th century and much of the 20th century, educational policies for indigenous groups in Mexico were viewed as tools for crafting a homogeneous national identity around the idea of mestizaje.[7] Public institutions aimed for the cultural assimilation of these population groups through schooling and teaching of Spanish language. Indigenous cultures during the time were relegated and, common to the times, approached as if inferior to European ones (Stavenhagen, 1988, Salmerón and Porras 2010).

Organizations of native populations appeared after the revolution (1910-1917). Many of them originated in the 1930s but they only gained strength in the early 1970s. It was in this later decade in which public education shifted to adopt a multicultural and multilingual approach. Federal resources destined for the National Indigenous Institute grew more than tenfold between 1971 and 1976 (Sarmiento 1985)\footnote{(from 39.1 mill in 1971 to 466 mill in 1976).} and in 1975 the first National Congress of Indigenous Populations was held.[8] The congress was the catalysing event for the creation of the National Council of Indigenous Populations [9] where, for the first time, representatives of indigenous groups would work together in a national political organization (Recondo 2007).

The creation of this national indigenous council, together with the debate about multiculturalism and education of the time, led a series of institutional changes. In particular, bilingual education became a goal in itself rather than a vehicle for cultural homogeneity (García Segura 2004, Jiménez and Mendoza 2015, Jiménez-Naranjo and Mendoza-Zuany 2016).[10]

In 1978, a reform established that education would be imparted in the mother tongue of the child at least during the first years of primary school. The new focus on education would initially look only at the linguistic component as a differentiator, relegating the cultural element. Implementation of the reform took time due to technical difficulties but in 1984 textbooks, programs, guides, learning material and general books in over 20 indigenous languages were produced (Salmerón and Porras 2010). This material was created for pre-schooling and the first four years of primary school.

The next set of reforms occurred as a result of political pressure during the 1990s, a period that also saw the EZLN uprising (an ideological and armed movement led by indigenous in the state of Chiapas in 1994). Among the most significant accomplishments in favour of indigenous groups was a reform

[7] The term "Mestizo" is a racial categorization from colonial times used to refer to a descent of a combined Spanish and American. This concept ignored the fact that within each region of the continent, now Latin America, independent cultures and civilization prevailed

[8] Occurred in Pátzcuaro, Michoacán.

[9] CNPI Spanish acronym

[10] This shift was driven by organizations such as the National Alliance of Bilingual Indigenous Professionals which was founded in 1977.

in Jan 1992 (Art 4)[11] recognising the constitutional right of indigenous communities to self-determination. The reform aimed to guarantee the right of these groups to preserve and enrich their languages, knowledge and culture. This reform would have important governance and administrative changes for indigenous communities long after.

In January 2001 the Federal Government created a national institute to coordinate bilingual and intercultural education (https://eib.sep.gob.mx/). This institute is in charge of developing educational curricula to attend cultural diversity, forming teachers, producing learning material, and pertinent school models. The Law of Linguistic Rights (2003), grants students of basic education the right to receive their education in their mother tongue, regardless of their location, a legal upgrade to the 1978 reform outlined above (Schmelkes 2006).

As can be seen, the institutional framework for the protection of indigenous languages is limited. Formal mechanisms for the protection of cultural identity need to be accompanied with resources for them to be effective. The policy mix of the kind seen in European nations is not really present in the context and, as a result, much of the transmission rely on informal mechanisms, mainly those of the household and the society in which individuals live.

The composition of the household and the characteristics of society are central in explaining the transmission of language. The role that each of these networks play is to some extent distinct. The easiness of learning a language will be a function of how the household is composed, that is of how many other indigenous language speakers there are in the household. As for societal networks, whether they are employment networks or of a more casual nature, these will be in the core of how valuable a language is. From an individual perspective, the value of a language is an increasing function of the number of other actual speakers. This is an example of spillover effects that is consistent with the existence of indigenous language clusters.

[11] In current Mexican Constitution, as a result of another constitutional reform in 2001, the changes of Article 4 in 1992 have been shifted to belong to Article 2.