Foreign language skills and labor market outcomes: The case of English in Mexico*

Oscar Galvez-Soriano†
University of Chicago

February 2024
(Latest version here)

Abstract

In this paper, I study the prevalence of English skills and the labor market returns to English skills in Mexico. I use individual-level data from the 2014 Subjective Well-being Survey, which, unlike other large nationally representative data sets, includes a measure of English proficiency. To address concerns regarding endogeneity in the relationship between English skills and labor market outcomes, I exploit policy changes in various Mexican states that introduced English instruction (as a subject) in public elementary schools during the 1990s. Using a Two-Way Fixed Effects specification, I estimate the effect of these state policies introducing English instruction. Subsequently, I utilize the variation in English skills resulting from these policy changes to estimate the causal effect of English skills on labor market outcomes. I offer robust estimates in the presence of heterogeneous treatment effects due to variations of the treatment over time and across treated regions. My findings indicate that these policies offering English instruction did not significantly affect wages, shifted workers out of physically demanding occupations, and increased the likelihood of speaking English.

JEL Classification: I21, I28, J24, J31.
Keywords: Primary School, Education Reform, Occupational Choice, Labor Income.

*I am deeply thankful to Aimee Chin for her guidance and comments throughout this research. I am also grateful for the valuable comments of Chinhui Juhn, Jessica Serrano, Jorge Pérez, and Julieta Osornio. This work was supported by Banco de México, the Mexican National Council for Science and Technology (CONACYT), the University of Houston, and the University of Chicago. I benefited from presenting at the University of Houston and Banco de México research seminars and the Southern Economic Association annual conference. This research also benefited from my Summer Fellowship at Banco de México in 2022.

†Kenneth C. Griffin Department of Economics. Chicago, IL 60637. e-mail: ogalvez@uchicago.edu

Introduction

Given the use of English as a *lingua franca* in the global economy and Mexico's proximity to the United States (US), there could be an economic value to English skills in Mexico. In this context, there has been an unprecedented expansion of English programs to offer English instruction in the public education system of non-English speaking countries. This expansion has been commonly motivated by the assumption that English instruction will enhance English language skills and, ultimately, improve labor market outcomes. Surprisingly, there is scarce research on the prevalence of English skills and labor market returns to English skills in Mexico. Indeed, the existing literature has traditionally studied the effect of English abilities on earnings in English-speaking countries, while there is limited research in contexts where English is not the predominant language of communication. The first evidence suggesting a positive correlation between English skills and earnings, in this context, was provided by Azam, Chin and Prakash (2013) with evidence from India and Eriksson (2014) with the case of South Africa.

Nevertheless, much remains unknown about the effectiveness of exposure to English instruction on the acquisition of English abilities and the role of these abilities on earnings. Angrist, Chin and Godoy (2008) offered the initial evidence that changing the medium of instruction (from English to Spanish) did not impact the English proficiency of Puerto Ricans. In contrast, Eriksson (2014) found a positive effect on English proficiency when the medium of instruction shifted from Afrikaans to English. However, the majority of non-English-speaking countries globally have introduced English as a subject rather than as the medium of instruction. For instance, the first study to exploit a policy change in exposure to English instruction (as a subject), finds a positive association between the probability of exposure to English instruction and wages (Chakraborty and Bakshi, 2016). However, they do not actually measure if this exposure leads to the acquisition of English skills. Likewise, although Gálvez-Soriano (2023) has a more robust measure of exposure (weekly hours of English instruction), he cannot provide causal evidence that exposure leads to the formation of English abilities due to a lack of necessary data. Instead, he offers evidence that exposure does not affect other cognitive skills, suggesting that his findings are consistent with the acquisition of English abilities.

Furthermore, evidence on the returns to English skills is scant in the context of non-English speaking countries and particularly rare in Latin American economies. Most existing research highlights the positive association between English language skills and immigrants' earnings in English-speaking countries (see Isphording (2014) and Chiswick and Miller (2015) for a review). This literature provides evidence that immigrants with English skills in the US have higher wages and better educational attainment. Similar results on wages have been found in the context of immigrants to Australia and European countries (Dustmann, 1994; Chiswick and Miller, 1995; Dustmann and Soest, 2001; Hayfron, 2001; Shields and Price, 2002; Williams, 2011). Likewise, it has been observed that the English premium among immigrants could vary depending on workers' age and education; younger and more educated immigrants tend to have greater returns to English at work (Lang and Siniver, 2009; Azam, Chin and Prakash, 2013). Nevertheless, our understanding of the effects of English instruction on the acquisition of English abilities and the returns of these skills in non-English speaking

countries is still limited.

To address these gaps in the literature, I explore the following research question: What are the returns to English skills in a non-English-speaking country? The context of my study is Mexico, an economy that primarily uses Spanish as its main language of communication. Furthermore, the significance of English in Mexico could be even more pronounced than in other developing countries due to its close relationship with the US in terms of trade and migration. Additionally, the external validity of my context is potentially broader than in the case of India or South Africa, which differ significantly from most developing countries as former British colonies where English abilities are often prevalent among the upper social class and government workers.

To answer my research question, I leverage individual-level data from the 2014 Subjective Well-being Survey, a dataset unique in including a measure of English proficiency, unlike other large nationally representative datasets. Specifically, respondents are queried about their English-speaking ability. Despite having a measure of English proficiency, estimating the causal effect of English skills remains challenging. The difference in labor market outcomes between individuals who speak English and those who do not cannot be interpreted as the causal effect of English skills due to potential selection biases in those who speak English. English-speaking ability is likely correlated with other variables influencing labor market outcomes. To estimate the causal effect, I exploit the implementation/expansion of six state English programs in Mexico. In my identification strategy, I employ locality-by-cohort variation in exposure to English instruction (driven by state policies since the early 1990s) within a staggered Difference-in-Differences (DiD) framework. Subsequently, I explore occupational choices to better understand the mechanisms behind the zero effect on wages.

Hence, in this paper, I provide the first empirical evidence that exposure to English instruction may lead to the formation of English skills in Mexico. The staggered DiD estimate from the first stage equation suggests that these English programs increase the likelihood of reporting speaking English by 11.2 percentage points. The average size of the intervention I study is relatively small. It increased exposure to English instruction, on average, by almost 28 minutes per week. Scaling this up, it would suggest that one additional hour of English instruction per week would increase the likelihood of speaking English by 24.3 percentage points.

Furthermore, I present some of the first causal estimates on the impact of English programs on labor market outcomes in the context of a non-English-speaking country. The set of labor market outcomes I examine includes the likelihood of working for pay, wages, and occupational choices. I find that the average intervention does not affect either the likelihood of working for pay or wages, although the point estimate for the latter is positive. Finally, I also provide evidence that programs offering English instruction at school did increase school enrollment. This finding suggests that the long-term effects of the English programs may positively influence wages through an increase in schooling.

There are three related papers in the context of Mexico. Delgado Helleseter (2020) offered the first estimate on the returns to English abilities in the market of jobs posted online. He finds that the wages of Mexican English speakers are approximately 28 percent higher than those of non-English speakers. However, his sample is composed only of online advertisements from a single online job board, which implies concerns about the sample selection and the

external validity of his results. Second, Charles-Leija and Torres (2022) estimate the returns to English skills in Mexico using the same data set that I use in this paper. However, they do not address the concern about the endogeneity of English skills beyond controlling for some observable characteristics. Finally, Gálvez-Soriano (2023) provides the first empirical evidence that exposure to English instruction does not affect the wages of the mean worker in the Mexican formal sector. However, he does find a positive and significant effect on wages only among high-achieving individuals.

The remaining of this paper proceeds as follows. In the first section, I offer a comprehensive description of the English abilities in Mexico along with the background of the policy changes I exploit in this paper. In section 2, I describe the database I use. In section 3, I explain the empirical strategy. In section 4, I first show the results of a naive estimate on the returns to English skills, then I offer an estimate of the effect of English programs on the acquisition of English abilities and labor market outcomes. In this same section, I also provide robustness checks and analyze the potential mechanisms. Finally, section 5 summarizes with a discussion of my findings and a brief conclusion.

1 English language skills in Mexico

1.1 Prevalence of English skills among the Mexican population

The descriptive analysis in this section relies on the English proficiency measure available in the 2014 Mexican Subjective Well-being Survey (BIARE, acronym in Spanish), detailed further in the data section. This measure is derived from survey respondents' answers to the question of whether they speak English. While not a detailed measure of English language skills, it remains a valid gauge of English proficiency, and this makes it feasible to describe the prevalence of these types of skills in Mexico and analyze their relationship with labor market outcomes. Using this measure of English proficiency, I document five empirical observations. First, there exists a gender-English gap, with more men declaring proficiency in English than women. Second, increased exposure to English instruction correlates with enhanced English skills. Third, English proficiency rises with educational attainment. Fourth, indigenous people lag in the process of learning English, and similarly, rural areas trail in this learning process.

BIARE surveys only adults aged 18 and older, but I focus on the range 18-65 to primarily capture individuals participating in the labor market (excluding retired workers). Additionally, I provide a comparison between Mexican states that implemented English programs in public primary schools and states that did not. Hence, this preliminary overview highlights two significant dimensions of heterogeneity in English abilities: among types of individuals and among geographical regions.

The primary language spoken in Mexico is Spanish, and all official documents are written in this language. As of 2014, approximately 7% of the Mexican adult population could speak English (see Table 1). Similarly, data from the 2020 population census suggests that 6.1% of Mexicans declared being able to speak an indigenous language. In this initial descriptive analysis, I observe a positive correlation between English instruction and English abilities in

Mexico. Through a regional comparison, I find that states implementing English programs in primary schools have a higher proportion of individuals with English skills than other states.[1] Thus, it could be inferred that exposure to English instruction may facilitate the acquisition of English skills.

I document the existence of a gender-English gap in Mexico, with more men declaring to speak English than women. About 9.5% of the male population in Mexico speaks English, while 5.1% of the female population does. Notice, however, that since the English skills variable is self-reported, it could potentially overstate the gender-English gap, with fewer women reporting having English abilities. On the other hand, as in the former finding (for the overall Mexican population), there are more men (and women) with English abilities in states that have offered English instruction in primary schools than in those that did not. This latter result motivates the research question of whether exposure to English instruction improves the acquisition of English abilities in the context of a non-English speaking country.

Young adults with more exposure to English instruction report higher English skills than older individuals. I derived this observation from two facts. First, younger individuals (18-35 years old) are more likely to speak English than elderly individuals (51-65 years old). This higher likelihood could be attributed to the English programs implemented in the early 1990s and early 2000s in Mexico, which predominantly affected younger birth cohorts (as explained in subsection 1.4). Secondly, particularly among the younger birth cohorts, significant differences in English abilities emerge when comparing individuals in states with English programs to those in states without such programs. Once again, this finding supports the hypothesis that English programs enhance the acquisition of English abilities.

The ability to speak English increases with educational attainment. The proportion of college graduates (and higher attainment individuals) with English-speaking skills in Mexico is more than three times the corresponding proportion for all Mexican adults. This proportion is about the same between individuals with upper secondary and the national figure. On the other hand, less than one percent of the individuals with incomplete primary school speak English. These results suggest that either most of the acquisition of English skills is held during higher education or that most individuals who can afford higher education are likely to learn English. Furthermore, this also points out the importance of including education in my empirical analysis as a non-linear variable. Indeed, in all models shown in section 3, I control for education fixed effects to capture the non-linear relationship between education and English skills.

I also document the existence of an ethnicity-English gap in Mexico, with nearly 7.3% of non-indigenous people with English skills and almost 2.5% of indigenous people who speak English. I claim that this is a result of the lack of English education in marginalized areas where indigenous people have settled. In fact, most of the efforts to teach indigenous people a second language are for them to learn Spanish than a foreign language. Another piece of evidence is the fact that I do not find differences in English abilities between indigenous people living in states that have implemented English programs and indigenous people living in states without these programs, which may suggest that the state English programs have

[1]States with English programs include Aguascalientes, Durango, Nuevo Leon, Sinaloa, Sonora, and Tamaulipas. While some other Mexican states have offered English instruction in public primary schools, it has been irregular, with only a few beneficiary schools and/or no expansion over time.

not reached the marginalized villages where indigenous people live.

Similarly, there is a considerable geographical variation in the prevalence of English skills in Mexico. In particular, I find a substantial difference in English ability between urban and rural areas, with the former having 8.3% of English speakers, while the latter have only 2.3%. Previous estimates by CIDAC (2008) suggested that this number was 6% for the urban adult population in Mexico. The story behind this difference between the urban and rural contexts is similar to the indigenous people case. Indeed, the state governments that have implemented English programs relegated rural and marginalized areas preventing their schools from implementing English programs. The reason is that most of the first English programs consisted of virtual courses because of the lack of English teachers, which means that schools without the equipment needed (computers, access to the Internet, projector, etc.) could not benefit from these programs. This issue has characterized the context of rural schools in Mexico.

1.2 Geographical heterogeneity of English skills

In this subsection, I present five observations concerning the geographical heterogeneity of English in Mexico. First, the population with English abilities is mainly located in six Mexican states: Aguascalientes, Baja California, Baja California Sur, Jalisco, Queretaro, and Quintana Roo. Second, the less industrialized Mexican states (South and Southeast regions) have a small share of English speakers. Third, English-speaking individuals are primarily found in urban areas. Fourth, the states with more exposure to English instruction are not necessarily those with more English speakers. Finally, administrative records suggest that six out of 32 Mexican states consistently offered English instruction in public primary schools during the early 1990s.

More than half of the Mexican states have shares of English-speaking individuals higher than the figure for the national level (6.99%), from which four states more than double this proportion (Aguascalientes, Baja California, Baja California Sur, and Quintana Roo). The Mexican states of Quintana Roo and Baja California Sur likely have more English speakers than most of the Mexican states because they are located in tourist regions with the two most famous destinations: Cancún and Los Cabos, respectively. On the other hand, Aguascalientes hosts international companies from the automotive assembly, auto parts manufacturing, industrial robotics, and electronics manufacturing, which are suppliers of Nissan, Mazda, Mercedes-Benz, GM, Honda, Infiniti, and Volkswagen, among others. Likewise, Baja California manufactures for international companies in the industries of aerospace, automotive, electronics, and medical device manufacturing. But also, because of its closeness with the US, some Mexican workers living in Tijuana commute every day to work in American companies located in Chula Vista and San Diego.

The South and Southeast region of Mexico has a low proportion of English speakers except for the state of Quintana Roo. This region is well known for its high proportion of people in poverty condition. In fact, the seven states colored in light gray from panel (c) of Figure 1 are classified among the poorest Mexican states according to the biannual report of CONEVAL (2020), except for the State of Mexico and Guanajuato, which are around the middle of the distribution (ranked in the place 12 and 17, respectively). Two potential reasons could

explain the lack of English speakers in this region. First, the poor quality of education, with these states having the lowest average grades in the ENLACE test. And, second, the lack of incentives due to the poor economic activity, with this region contributing the least to the national economic activity (except for Tabasco, which contributes substantially to the sector with oil extraction).

Mexican rural areas exhibit the lowest shares of English speakers. However, there are seven with relatively higher shares: Aguascalientes, Baja California, Baja California Sur, Durango, Hidalgo, Nayarit, and Zacatecas (refer to panel (d) of Figure 1). The first three states have a relatively higher share of English speakers due to the reasons mentioned in observation number one. Durango and Zacatecas serve as international migrant sender states, particularly to the US, with a relatively greater proportion of migrants in rural areas compared to urban ones. Additionally, the state of Hidalgo hosts international companies in automotive, auto parts, technology, telecommunications, and transportation such as GEMI International, Motorola Solutions, Grupo Marpa, and Transtell, situated in rural regions of the state. Meanwhile, Nayarit is renowned for exporting agricultural products (figs, pineapples, avocados, guava, mangoes, cantaloupes, watermelons, and papayas) to the US, which may explain the non-negligible share of English speakers in the rural context.

The Mexican states with more exposure to English instruction are not necessarily those with more English speakers. Indeed, most of the states colored in darker gray in panel (c) of Figure 1, reflecting a high share of English speakers, are touristic, migrant senders, and/or hosts of international companies, as explained before. For states that coincide in terms of exposure and the proportion of English speakers (Aguascalientes, Sonora, Morelos, Nuevo Leon, and Queretaro), it is difficult to disentangle the reasons for this high proportion, but I will explore some of them later in this paper. As for the remaining states (Quintana Roo, Baja California, Baja California Sur, Nayarit, Chihuahua, and San Luis Potosi), we could infer that the formation of English abilities is not directly explained by the exposure to English instruction.

Finally, administrative records suggest that six out of 32 Mexican states offered English instruction in public primary schools during the early 1990s. Indeed, using historical administrative data from the Mexican school census, I find that the states of Aguascalientes, Durango, Nuevo Leon, Sinaloa, Sonora, and Tamaulipas have implemented state English programs, which increased the weekly hours of English instruction offered in Mexican public schools (see Figure A.1 from the Appendix).

1.3 Occupations and English skills in Mexico

In this subsection, I describe the economic occupations in which Mexican English speakers concentrate. I document three main observations. First, elementary and manual unskilled occupations do not "require" English abilities in Mexico. Second, clerical support, professionals, and managerial occupations concentrate most English speakers in Mexico. Third, occupations with more English speakers pay more to their workers, have a more even proportion of females, and these workers are more educated.

The occupations that do not require English abilities in Mexico are farming, elementary,

and crafts occupations (see the first categories in panel (a) of Figure 2).[2] This finding is not surprising because these types of occupations do not require a high level of communication skills and, instead, these occupations are more manual-intensive. Thus, on average, in these occupations, only two out of a hundred workers speak English. The English speakers in these occupations likely acquired their English abilities either in school or through a migration network. On the other hand, these occupations are among the worst paid. This finding suggests a positive correlation between English skills and earnings. Notice, however, that individuals working as machine operators also have a low proportion of English speakers and still earn as much as clerical workers (who have a high proportion of English speakers). This exception will be particularly relevant in terms of occupational decisions for individuals exposed to English instruction, as I explain in subsection 4.3.

On the other hand, clerical support, professionals, and managerial occupations concentrate most of the English speakers in Mexico (see the last four categories in panel (a) of Figure 2). In the case of clerical support, it is natural to have more English speakers since these types of occupations are intensive in communication skills. However, it is also true that most of these occupations, for example, assistants and secretariats are not required to have English abilities, except for executive positions or in tourism and international companies. On the other hand, professionals and technicians are more likely to have English abilities because, in upper secondary and professional education, the English language subject is compulsory and required to graduate with some degrees. Managerial occupations are the most likely to require English skills, with three times the national average proportion of English speakers. This latter finding is consistent with tasks requiring communication skills and public relations, especially in medium-sized and large companies. Finally, 40% of the individuals working abroad have English skills, consistent with Mexicans who cross the US border every day to work in American companies. The remaining occupations (customer service and sales occupations) are below the average English skills in Mexico.

Among the occupations with more English speakers, workers are better paid, more educated, and there is a more even proportion of females. The average monthly wage in Mexico is 6,262 pesos (for adults 18-65 years old in 2014 who work for pay), while monthly wages in clerical support occupations are about 1.1 times higher; professionals and technicians earn 1.7 times the mean wage in Mexico, while the mean wages in managerial occupations is 2.4 times higher. This result suggests a positive correlation between English abilities and earnings. Likewise, as previously noted in subsection 1.1, I find a positive correlation between English abilities and education. Finally, customer service, sales, and clerical support occupations have the most even proportion of female workers.

1.4 Policy change: English programs in Mexican states

Since the early 1990s, several Mexican states implemented English programs to offer English instruction in public primary schools. An important motivation for these states was the recently signed North American Free Trade Agreement (NAFTA) in December 1992, which

[2]Notice, however, that it is strong to claim anything about English requirements because the proportion of English speakers is not the same as the demand for workers with English abilities, but there should be a high correspondence.

came into force on January 1, 1994. In particular, English instruction in public primary schools is meant to improve the acquisition of English skills to facilitate labor mobility from Mexico to the US and Canada. Before launching these state English programs, only private schools offered English instruction, leaving more than 90% of the population unattended, which could potentially benefit from NAFTA. As previously noted in subsection 1.3, using administrative data from the Mexican school census, I identified six states that have implemented/expanded English programs in public primary schools, consistently increasing the hours of English instruction offered in these schools. Hence, in this section, I will describe how the implementation occurred in most of these states. However, some of these states do not have a publicly available registry of their English programs, implying that my identification strategy would be mainly data-driven for those particular cases.

The first two Mexican states that offered English instruction in primary schools were Nuevo Leon and Sonora. The English program in the Mexican state of Nuevo Leon was launched in 1993, initially benefiting only high-achieving students of fourth, fifth, and sixth grades in one hundred randomly chosen elementary schools.[3] However, it was not until the year 1998 that the state English program implemented English instruction as part of the regular curricula in the participant schools, and only in sixth grade. This expansion is one of the focuses of exploration in this paper. Over time, the state government increased the English program's coverage among elementary schools, including some preschools. For instance, in 2008, the program expanded coverage to students in fifth grade in schools already benefiting from the program in sixth grade. By 2008, the state English program covered approximately 60% of all elementary schools in Nuevo Leon.[4]

The state of Sonora also launched its English program in the year 1993 as a trial stage, but it was not until the year 2004 that the English language was incorporated as a subject in the regular curricula of public primary schools. In the expansion of 2004, the program aimed to offer English instruction to only the first and second grades of 10 out of 72 counties. After that, the program gradually expanded the coverage to all grades that comprise primary school in Mexico (from first to sixth) and to more beneficiary counties, reaching a state coverage of almost 50% by the year 2010 (Reyes Cruz, Murrieta Loyo and Hernández Méndez, 2011).

Then, the state of Tamaulipas launched an English program in the 2001-2002 school year, initially offering English instruction to the fourth grade of urban primary schools in the state, benefiting 44,777 students. For the 2003-2004 school year, the state English program expanded its coverage from fourth to sixth grade. Subsequently, in 2005, the program expanded to cover all six grades that comprise primary school in Mexico. In the 2011-2012 school year, all preschools in the state incorporated into the English program.[5] The Tamaulipas policy change I exploit in this paper is the first implementation of 2001.

As for the remaining states—Aguascalientes, Durango, and Sinaloa—no official sources indicate the process of the implementation of their English programs. However, administrative data from the Mexican school census suggests that these states launched their English

[3]Elementary schools in Mexico comprise from first to sixth grade, middle schools comprise from seventh to ninth grade, while high school comprises from tenth to twelfth grade. All three educational levels are part of the basic compulsory education system.

[4]All this information is publicly available on the Nuevo Leon English program's website.

[5]All this information is publicly available on the Tamaulipas English program's website.

programs in 2001, 2002, and 2004, respectively (see Figure A.1). Unofficial sources point out that the implementation of the English program in Durango was in 2002 as a pilot program, which operated for six years in a few public primary schools. In 2008, English instruction was incorporated into the regular curricula of the beneficiary primary schools. In 2009, the program reached 20% coverage of the students enrolled in Durango primary schools.

2 Data

My primary source of information is the 2014 Mexican Subjective Well-being Survey (BIARE, for its acronym in Spanish). BIARE is a representative survey of the Mexican population at the national and state levels. In 2014, this survey was conducted as part of the Mexican Household Income and Expenditure Survey (ENIGH) and served as an annex to the Socioeconomic Conditions Module. This arrangement allows for variables characterizing the socioeconomic conditions of individuals and households to coexist in the same database, currently employed to measure poverty in Mexico.

The 2014 BIARE round is notable as it included a one-time inquiry into individuals' English-speaking abilities. Respondents of this survey are adults aged 18 years and older, with the survey focusing on one adult per household. Consequently, my sample size is smaller than that in the ENIGH 2014 for the same age range. BIARE survey interviews are all conducted face-to-face, requiring information to be provided exclusively by the respondent, not through a third party. Respondents are asked a set of questions concerning their demographic and economic characteristics, perceived well-being, and their ability to speak English.

The second source of information I utilize is the Mexican school census, also known as Statistics 911. This census enables the identification of public schools that have offered English instruction in Mexico. This information is crucial for constructing a variable representing exposure to English instruction, allowing me to gauge the extent of state English programs implemented in Mexico since the early 1990s. To construct this exposure variable, I exclusively consider public elementary schools in the morning shift, as afternoon shifts are unstable, with the same school potentially not providing both shifts every academic year.

I measure 'exposure', by cohort and by locality, as weekly hours of English instruction using the ratio of total weekly hours of English instruction in each school to the total number of classes.[6] For each school-cohort, I calculate the average hours over the six years comprising primary school in Mexico. Finally, I determine the average, by cohort, of all schools in a given locality, weighting by the number of students per school. Hence, the locality average is always smaller than the figure per school because most primary schools in Mexico do not offer English instruction.

I link my exposure variable to the BIARE database by cohort and locality. In the BIARE database, the birth cohort of 1996 attended sixth grade (the final grade of primary school in Mexico) in 2007. Hence, the data I constructed for 2007 using the Mexican school census

[6] A 'locality' is the smallest geographically delimited area in Mexico, ranging from a single household in a rural context (*rancheria*) to over 100,000 inhabitants in larger Mexican cities. The next geographical size is the county (municipality), composed of several localities, and states, which are further subdivided by these counties.

includes the average exposure from 2002 to 2007 for the cohort 1996. Since I impute the average exposure at the locality level, my estimate of the policy change on exposure could be understated because this exposure variable includes schools with zero hours of English instruction. Nevertheless, this exposure measure remains informative. Furthermore, the final database I built in this paper is the first (in Mexico) to include both variables: exposure to English instruction and English skills.

I also explore the effect of English programs on occupational choices as a potential mechanism that mediates the impact on wages. In this part of the analysis, I use the O*NET classification of occupations as it provides detailed information on the tasks required by each job. I started using the 2011 Mexican System of Classification of Occupations (SINCO, for its acronym in Spanish), as it is the official classification of occupations in Mexico. Then, I crosswalk the 2011 SINCO with the 2010 Standard Occupational Classification (SOC) that is officially used in the US. The latest O*NET classification provides the crosswalk with the 2018 SOC, so I also worked on the crosswalk of the 2010 SOC and 2018 SOC. Finally, I used the O*NET classification by work activities. In particular, to define physically demanding jobs, I averaged the following two classifications: (i) Handling and Moving Objects, and (ii) Performing General Physical Activities. The other classification I use in this paper is Communicating and Interacting, which I consider as a proxy for jobs requiring communication skills.

The final database enables the identification of individuals with English abilities and the hours of English instruction to which they were exposed in primary school. I only consider respondents born between 1981 and 1996 who work for pay. These birth cohorts correspond to the youngest individuals observed in the BIARE survey, including those who may have had exposure to English instruction (younger cohorts) and those who did not have exposure (older cohorts).

A preliminary analysis between Mexican English speakers and non-English speakers (see Table 2) suggests that the former are more educated probably because they are acquiring English abilities in school; they earn more, which is consistent with the generalized idea of the positive returns to English skills in non-English speaking countries; and most of the English speakers are located in urban areas. On the other hand, Mexican English speakers are younger, potentially because they were more likely to have exposure to the recently implemented English programs in Mexico. Finally, I also show that English speakers are less likely to be women, indigenous, or married.

The descriptive analysis suggests that individuals with greater exposure to English instruction are more likely to possess English abilities (see the second row of Table 2). This aligns with my earlier analysis in section 1, where I showed that younger individuals (who are more likely to have exposure to English instruction) exhibit higher levels of English skills compared to older individuals. Furthermore, I showed significant discrepancies in English abilities between individuals residing in states with English programs and those in states without such programs. Finally, the occupation variables are dummies indicating a value of one if the "importance" of the required skill falls within the top quartile of the occupation distribution. For instance, according to O*NET, economists are not required to engage in physically demanding work or possess communication skills, while lawyers are required to have communication skills. The preliminary descriptive analysis suggests that English

speakers are more inclined to work in occupations necessitating communication skills and less inclined to work in physically demanding roles compared to non-English speakers.

3 Empirical strategy

We can model the relationship between English skills, Eng_i, and earnings, ω_i, using the following equation:

$$\omega_{isc} = \alpha + \beta\, Eng_{isc} + \boldsymbol{X_{isc}\Pi} + \epsilon_{isc}, \tag{1}$$

where ω_{isc} is the log of wages, the English skills variable, Eng_{isc}, is binary; it takes the value of one if the individual i speaks English and zero otherwise. I also include a vector of controls, $\boldsymbol{X_{isc}}$, with socio-demographic characteristics, such as education, gender, marital status, ethnicity, cohort fixed effects (FE), and locality FE.

However, the English skills variable could be endogenous in this wage equation. Two potential sources of endogeneity are: omitted variables and measurement error. First, the omitted variables issue arises from not controlling for unobservable individual characteristics such as abilities, which could be correlated with both English skills and wages. Second, it is likely that my English skills variable has measurement error as it captures self-reported ability. In this context, an OLS estimation would produce a biased estimate of β. Hence, instead of exploring the returns to English skills, I study the effect of English policies on labor market outcomes.

A staggered DiD specification, on the other hand, will allow me to offer an estimate of the causal effect of the English policies on labor market outcomes, y_{ics}. Let us define $HadPolicy_{sc}$ as the main effect variable, a dummy that takes the value of one if the individual i lives in a treatment locality and belongs to one of the affected cohorts, while it takes the value of zero otherwise. The reduced form equation is as follows:

$$y_{isc} = \theta + \psi\, HadPolicy_{sc} + \delta_s + \kappa_c + \boldsymbol{X_{isc}\Psi} + \varepsilon_{isc} \tag{2}$$

where ψ measures the effect of English programs on the acquisition of English abilities. In this specification, I fully control for state fixed effects, δ_s, common cohort effects, κ_c, and a vector of controls, $\boldsymbol{X_{isc}}$, with demographic and household characteristics as previously defined. Hence, these specifications could also be categorized as a Two-Way Fixed Effects (TWFE) model.

The consistency of my results will depend on the validity of the parallel trend assumption, which suggests that the change in the outcomes of interest between pre-treatment and post-treatment cohorts would have been the same in the treatment and the comparison states had the English program not been introduced/expanded in the former. I provide evidence that supports the validity of this assumption by analyzing the following event study-type equation:

$$y_{isc} = \theta + \sum_k \psi_{c-c_s^*}\, I_{(k=c-c_s^*)} + \delta_s + \kappa_c + \boldsymbol{X_{isc}\Psi} + \varepsilon_{isc},$$

where c_s^* denotes the first cohort affected by the intervention in state s. The difference $c - c_s^*$ is the time relative to c_s^*, with negative values reflecting older cohorts not exposed to the policy. $I_{(k=c-c_s^*)}$ is a dummy variable for $k = c - c_s^*$, so $\psi_{c-c_s^*}$ gives the effect of leads and lags of policy adoption. The omitted category is -1. Negative categories with zero effect validate the PTA, as shown in Figure 3.

4 Results

4.1 Descriptive analysis on the returns to English skills in Mexico

A descriptive analysis using a simple ordinary least squares estimation suggests that the returns to English skills in Mexico are zero, which contradicts the findings in the existing literature. In this part of the analysis, I use Equation 1, which controls for observable characteristics, cohort, and locality FE. In particular, I offer five estimates that progressively address the omitted variables problem by including more controls: first, a naive estimate (without controls); second, including cohort FE, gender, and ethnicity; third, including education FE; fourth, incorporating rural and marital status dummy variables; and finally, including locality FE. I also present estimates using low-education and high-education samples. The former isolates the effect of English abilities on wages from spillover effects of education, experience, and unobserved abilities. The latter includes the complementarity between English and education. Two main findings emerge from this initial approach: first, the omitted variables problem is primarily due to education, and second, English speakers do not earn more than non-English speakers.

Education accounts for approximately 89% of the selection bias arising from the observable omitted variables problem. A naive estimate would suggest that English speakers earn 124% more than non-English speakers in Mexico (0.807 natural log points). However, this estimate contains confounding factors such as gender and age. Once I control for these variables, the estimated effect reduces considerably to 94%. Furthermore, the most significant factor contributing to the omitted variables problem is education, reducing the estimate to 9.2%, though it is not statistically significant. Other crucial controls include socio-demographic characteristics. For example, individuals living in urban areas, and non-indigenous are more likely to speak English. Thus, I further mitigate the omitted variables problem by controlling for these characteristics, as shown in column (4) of Table 3. With this specification, the bias is further reduced, resulting in a negative point estimate that remains statistically insignificant. The remaining estimate (from column 5) mitigates the downward bias by accounting for non-time-varying unobservable characteristics at the locality level. These unobservables may capture characteristics of the poor localities that, after the English programs, make them more likely to offer English instruction, but that ultimately explain wage disparities.

Once I control for observable and unobservable characteristics of the individuals and their socioeconomic context, I find that English speakers do not earn more than non-English speakers. This conclusion derives from my proposed model, which implies two main assumptions. First, the set of controls I include in the model accounts for most of the omitted variables. Second, the remaining unobservable factors that affect the acquisition of English abilities are

constant over time and aggregated at the locality level. If we are willing to believe that these assumptions are valid, we may conclude that English speakers earn approximately the same as non-English speakers (see column 5 of Table 3). This result is different from previous findings of Delgado Helleseter (2020) and Charles-Leija and Torres (2022), who suggest that the wage premium of Mexican English speakers is 28% and 19.4%, respectively. Finally, I do not find significant returns to English skills among low-educated individuals.

4.2 Effect of offering English instruction and robustness checks

The previous estimates, however, do not provide a causal effect of English abilities on wages. In an attempt to provide a more reliable estimate, I propose a DiD strategy where I use locality-by-cohort variation in exposure to English instruction in primary school. With this strategy, I estimate the intention to treat (ITT) effect of offering English instruction in primary school. I study three main outcomes: exposure to English instruction, acquisition of English abilities, and the log of wages.

Panel A of Table 4 presents the results obtained from a traditional staggered DiD specification. First, the zero effect on the likelihood of working for pay (column (4)) rules out issues of sample selection. Second, I do not find a significant effect on wages. However, surprisingly, the point estimate is negative. Third, English programs may lead to the acquisition of English abilities. Indeed, English programs increase the likelihood of speaking English by 8.2 percentage points (see column (2) of Table 4). Furthermore, considering the hours of English instruction as an exogenous variable (due to the implementation/expansion of the English programs), we may argue that English programs in Mexico may have increased the likelihood of speaking English by 15 percentage points (0.082/0.546).

However, recent critiques in the DiD literature suggest that my estimates could be biased due to the presence of heterogeneous treatment effects (see de Chaisemartin and D'Haultfoeuille (2022) for a review). Indeed, the timing of the English programs varied among the treatment states, and schools that adopted these programs offered different amounts of hours of English instruction to different cohorts. Hence, in panels B and C, I provide robust estimates as proposed by Sun and Abraham (2021) and Callaway and Sant'Anna (2021). The results shown in Table 4 suggest that my original estimates for exposure and English skill acquisition are robust in terms of sign and significance and very close in magnitude to changes in the specification, which mitigates the potential bias due to heterogeneity in treatment effects across cohorts and over time.

In addition, I address the potential concern that the cohort span studied in this paper is "too wide" by narrowing it down. This concern arises because older cohorts may experience different job opportunities than younger ones, which could question the validity of the parallel trend assumption. Hence, I narrowed the cohort span to individuals born between 1985 and 1995. Panel D of Table 4 presents the results of this robustness check, demonstrating that my estimates are robust to changes in the cohort span.

Notice that the three different specifications, shown in panels A-C of Table 4, are inconclusive regarding the direction of the effect of English programs on wages. Hence, my best conclusion is that the effect was, on average, zero. Nevertheless, this result raises questions about the reasons for the lack of effect on wages. First, I analyze the effect on labor supply,

measured as the logarithm of hours worked per week (see column (5) of Table 4). Although the effect is statistically insignificant, the point estimate is consistently negative. Could this suggest an improvement in labor conditions? Second, column (6) of Table 4 may support this idea. In fact, I find that the English programs increased the probability of working in a formal job. However, although robust, this estimate is only significant at a 90% confidence level.

4.3 Mechanisms

In general, this intervention, which introduced English instruction in public elementary schools, could affect individuals' wages through three main channels: effects on individuals' abilities (including English skills), occupational choices, and educational attainment or school enrollment. I provide suggestive evidence that the zero effect on wages is not a result of trade-offs between English and other cognitive abilities but a result of occupational choices. Furthermore, it may be the case that once all individuals in the youngest cohorts are incorporated into the labor market, the effect on wages becomes positive and statistically different from zero.

First, it is natural to think that wages will vary along with changes in the marginal product of labor. For example, if the inclusion of English as a subject affected the teaching time of other subjects, there may be a trade-off between foreign language and other cognitive skills. This potential trade-off would produce an ambiguous effect on the marginal product of labor and, ultimately, on wages. However, the existing literature has shown that exposure to English instruction does not affect cognitive skills in Mexico (see Gálvez-Soriano (2023)). Furthermore, results from Table 4 provide strong evidence that the English programs increased the likelihood that individuals speak English. In light of these results, the increase in marginal product of labor was not accompanied with a significant increase in wages.

Second, another factor likely explaining the zero effect on wages could be related to the potential increase in the probability of having a formal job (as shown in column (6) of Table 4). In other words, English programs may have affected individuals' occupational decisions. Specifically, we are interested in analyzing if the state English programs improved occupational conditions. I utilize the O*Net classification of occupations, which assigns a score to each occupation based on the intensity of specific tasks required for that job. I focus on jobs that require physical work or those that demand communication skills, as these are likely to be the job characteristics individuals with English skills consider changing when making their occupational choices.

Indeed, the lack of effect on wages could potentially be explained by a trade-off individuals face between seeking better-paid jobs or improved working conditions. Point estimates indicate that the English programs increased the likelihood that individuals work in occupations requiring communication skills and are less likely to work in physically demanding occupations (see panels (a) and (c) of Figure 4). These findings are derived from robust estimates under heterogeneous effects, as suggested by Callaway and Sant'Anna (2021). Furthermore, these effects are driven by individuals with high educational attainment, which suggests a complementarity between education and English skills (see panels (b) and (d) of Figure 4). However, it's important to note that these results are not statistically signifi-

cant at the conventional level, yet the direction of the effects aligns with the aforementioned explanation.

Finally, the average effect of English programs on wages excludes potentially high earners who are currently enrolled in school. In fact, the youngest cohorts, who are more likely to be enrolled in school, have greater exposure to English programs. However, even when comparing within the same cohort, individuals affected by the program are more likely to be enrolled in school (see Figure 5). These are good news as we may expect a future improvement in human capital accumulation due to exposure to English programs, which has not yet been reflected in wages.

5 Conclusions

In this paper, I offer a description of English abilities in Mexico, which had not been done before due to the lack of data on English skills. Furthermore, I provide the first estimate of the causal effect of English programs on the acquisition of English abilities and wages in Mexico. This latter finding contributes to the existing literature, which has mostly offered evidence on the returns to English skills in the context of immigrants in English-speaking countries or former British colonies (India and South Africa). My results provide some of the first estimates in the context of a non-English-speaking country and the first among Latin American economies.

Using ordinary least squares estimation and controlling for a rich set of variables including education, I find that there is no significant correlation between English skills and wages. Additionally, the omitted variables problem, which produces a biased estimate, is mostly due to education. Furthermore, point estimates suggest that male English speakers earn more than female English speakers. Similarly, highly educated English speakers earn more than low-educated individuals. However, we may still be concerned about an endogeneity problem due to omitted variables such as abilities. Furthermore, there may also be a bias in my estimates due to a measurement error in the English skills variable.

To provide an estimate with a causal interpretation, I exploit the implementation of several state English programs that offered English instruction in public primary schools. In particular, I use a staggered DiD specification. My results suggest that the implementation of English programs does not have any effect on wages. However, the point estimate suggests a potentially positive effect. This zero effect on wages is not a result of changes in the accumulation of human capital through other cognitive skills (Gálvez-Soriano, 2023), but potentially through the acquisition of English skills. In fact, if I consider hours of English instruction as an exogenous variable (due to the implementation/expansion of the state English programs), we may argue that English instruction in Mexico could increase the acquisition of English skills by 15 percentage points (0.082/0.546).

Regarding the mechanisms, I find that individuals who had exposure to the intervention are more likely to have a formal job. This mechanism is consistent with a reduction in the likelihood of working in physically demanding jobs and an increase in the likelihood of working in occupations that require communication skills. These results suggest that although English instruction may not affect wages, it may expand employment opportunities. A re-

vealed preference argument suggests they cannot be worse off, as they could still choose the traditionally manual occupations if they wished. Finally, in the analysis, I am systematically excluding potentially high earners.

References

Angrist, Joshua, Aimee Chin, and Ricardo Godoy. 2008. "Is Spanish-only schooling responsible for the Puerto Rican language gap?" *Journal of Development Economics*, 85(1-2): 105–128.

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.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

Chakraborty, Tanika, and Shilpi Kapur Bakshi. 2016. "English language premium: Evidence from a policy experiment in India." *Economics of Education Review*, 50: 1–16.

Charles-Leija, Humberto, and Aldo Josafat Torres. 2022. "English Proficiency and Salary in Mexico." *Analisis Económico*, 30(94): 167–180.

Chiswick, Barry R, and Paul W Miller. 1995. "The endogeneity between language and earnings: International analyses." *Journal of labor economics*, 13(2): 246–288.

Chiswick, Barry R., and Paul W. Miller. 2015. "International Migration and the Economics of Language." In *Handbook of the Economics of International Migration*. Vol. 1 of *Handbook of the Economics of International Migration*, , ed. Barry R. Chiswick and Paul W. Miller, 211–269. North-Holland.

CIDAC. 2008. *CIDAC survey about human capital in Mexico.* Mexico City:Centro de Investigacion para el Desarrollo, A.C.

de Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2022. "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey." *The Econometrics Journal*, 26(3): C1–C30.

Delgado Helleseter, Miguel Antonio. 2020. "English Skills and Wages in a Non-English Speaking Country: Findings from Online Advertisements in Mexico." *The International Journal of Interdisciplinary Global Studies*, 15(3): 13–27.

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

Dustmann, Christian, and Arthur van Soest. 2001. "Language fluency and earnings: Estimation with misclassified language indicators." *Review of Economics and Statistics*, 83(4): 663–674.

Eriksson, Katherine. 2014. "Does the language of instruction in primary school affect later labour market outcomes? Evidence from South Africa." *Economic History of Developing Regions*, 29(2): 311–335.

Gálvez-Soriano, Oscar de Jesus. 2023. "Effects of English instruction and English skills on labor market outcomes in Mexico." PhD diss. University of Houston.

Hayfron, John E. 2001. "Language training, language proficiency and earnings of immigrants in Norway." *Applied Economics*, 33(15): 1971–1979.

Isphording, Ingo E. 2014. "Language and labor market success." *IZA Discussion Paper*.

Lang, Kevin, and Erez Siniver. 2009. "The return to English in a non-English speaking country: Russian immigrants and native Israelis in Israel." *The BE Journal of Economic Analysis & Policy*, 9(1).

Reyes Cruz, María del Rosario, Griselda Murrieta Loyo, and Edith Hernández Méndez. 2011. "Políticas lingüísticas nacionales e internacionales sobre la enseñanza del inglés en escuelas primarias." *Revista pueblos y fronteras digital*, 6(12): 167–197.

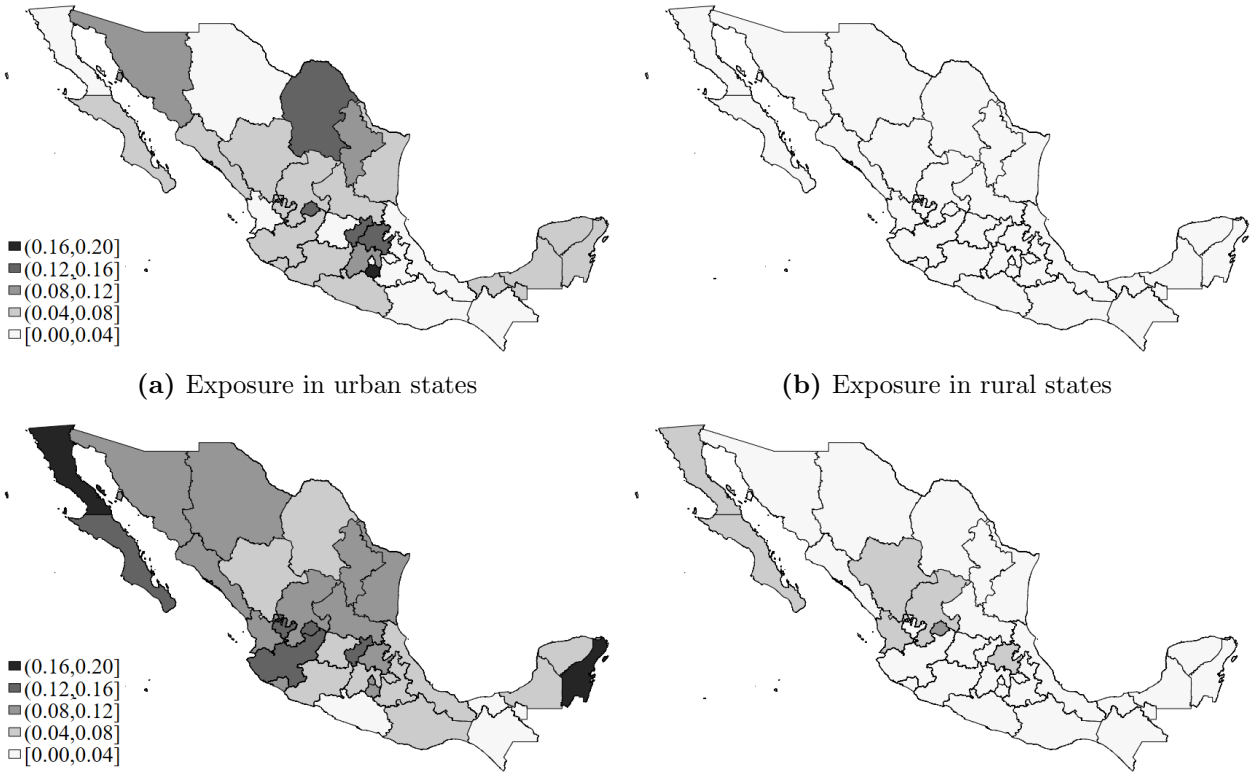
Shields, Michael A, and Stephen Wheatley Price. 2002. "The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas." *Journal of population Economics*, 15(1): 137–160.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Williams, Donald R. 2011. "Multiple language usage and earnings in Western Europe." *International Journal of Manpower*.

Figures and tables

Figure 1: Exposure to English instruction and English abilities in Mexican states



(a) Exposure in urban states

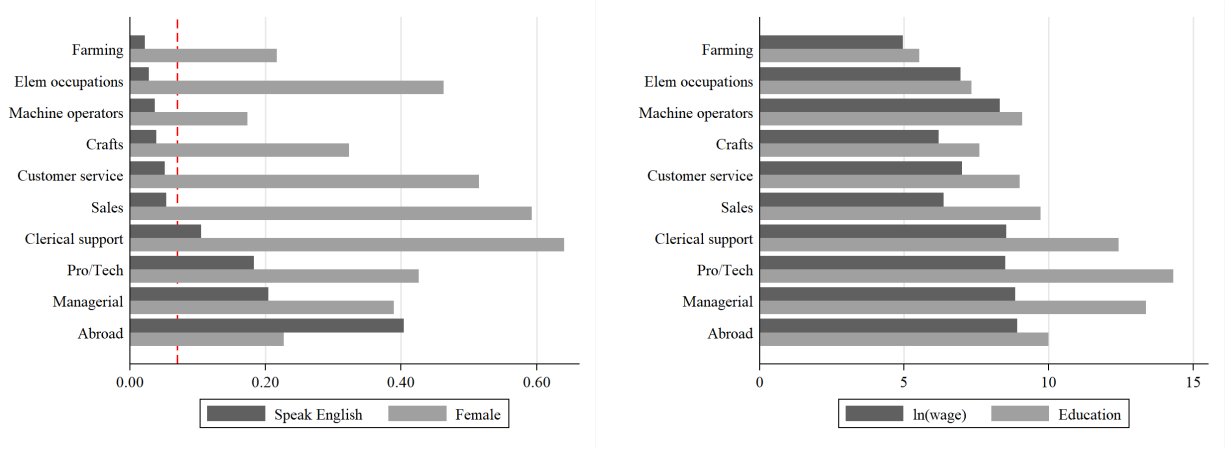
(b) Exposure in rural states

(c) English abilities (urban)

(d) English abilities (rural)

Note: Maps in upper panels (a) and (b) represent the proportion of Mexican public elementary schools that offered English instruction in a given state, in 2008. Maps in lower panels (c) and (d) represent the proportion of individuals aged 16–65 who self-reported their ability to speak English, using data from the 2014 Mexican Subjective Well-being Survey (BIARE).

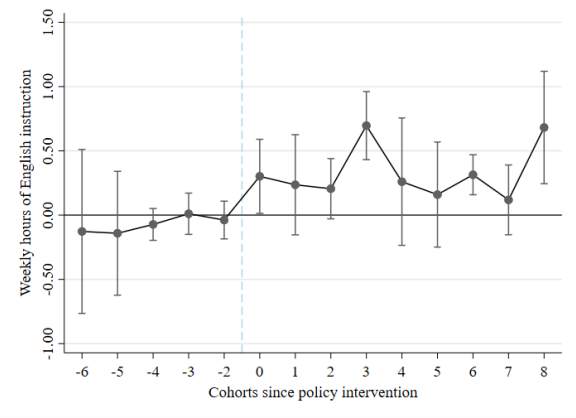
Figure 2: English abilities, wages and education by occupations



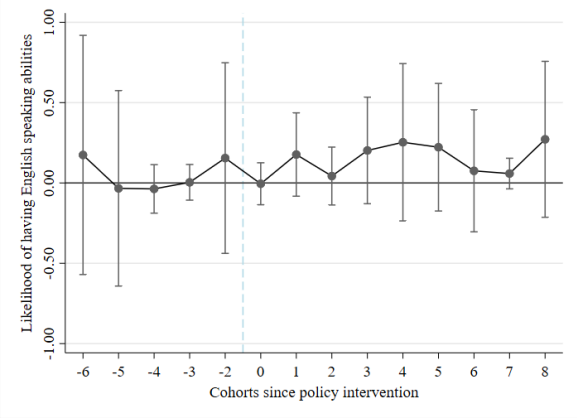
(a) Proportion of female and English speakers **(b)** Wages and education

Note: This figure shows the percentage of individuals who have English speaking abilities, mean wages, percentage of women and their mean schooling, by aggregates of occupations in Mexico. The occupations were determined using the International Standard Classification of Occupations (ISCO-08) from the International Labor Organization (ILO) at one digit code level. The managerial category includes supervisors from other occupations. The 'abroad' category contains individuals who reported working abroad, but it is a mix of all other categories. The sample contains Mexicans ages 18–65 who self-reported their ability to speak in English. The vertical dotted line represents the mean of English speakers in this sample 0.0699 (6.99%).

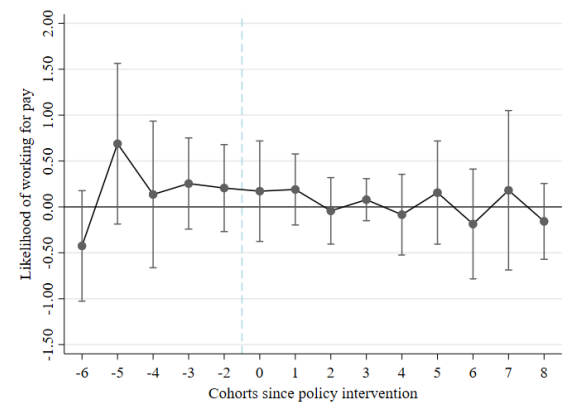
Figure 3: Event-study graphs from staggered DiD specification



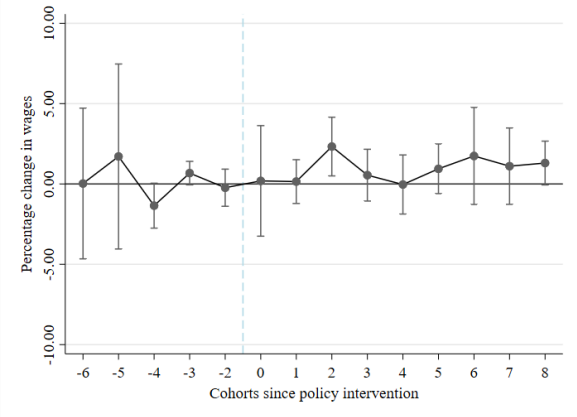
(a) Hours of English



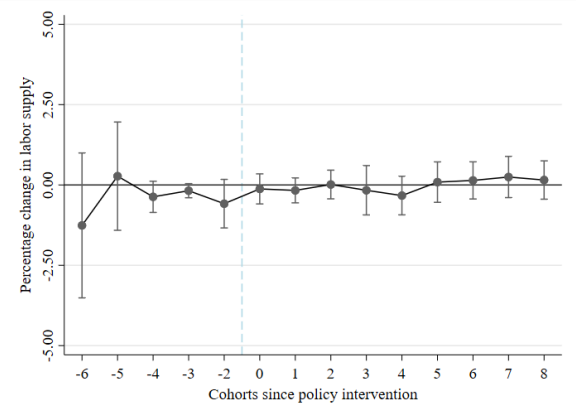
(b) Speak English



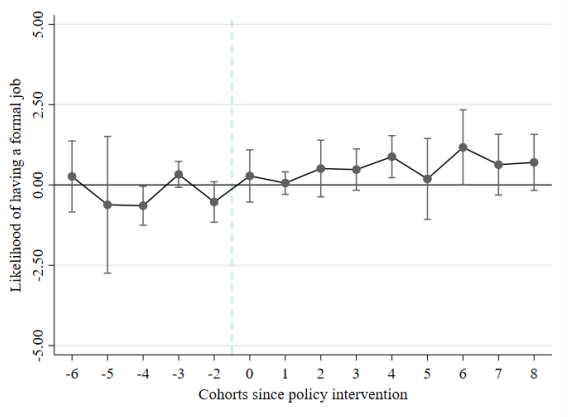
(c) Paid work



(d) Ln(wage)



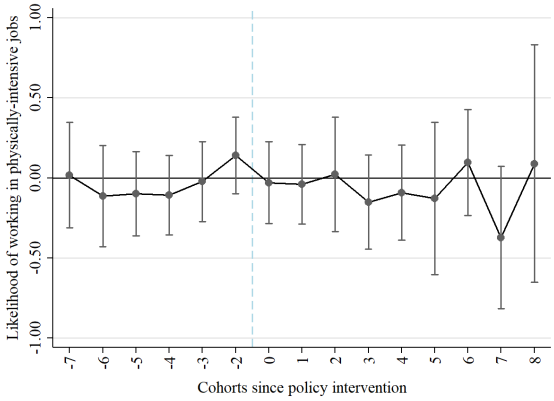
(e) Labor supply



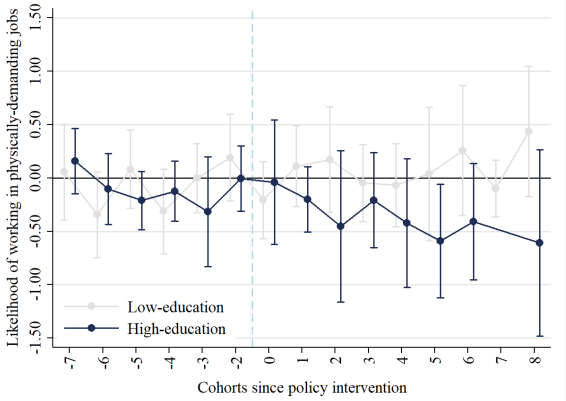
(f) Formal job

Note: This figure plots robust estimates (as proposed by Callaway and Sant'Anna (2021)) from the interaction terms between the treatment variable and an indicator function for each cohort since the policy intervention, in an event study type regression. The omitted period is -1. The vertical dotted lines indicate the moment of the intervention. The no statistically significant estimates at the left of the vertical dotted line suggest parallel trends before the policy. Confidence intervals at 95% level.

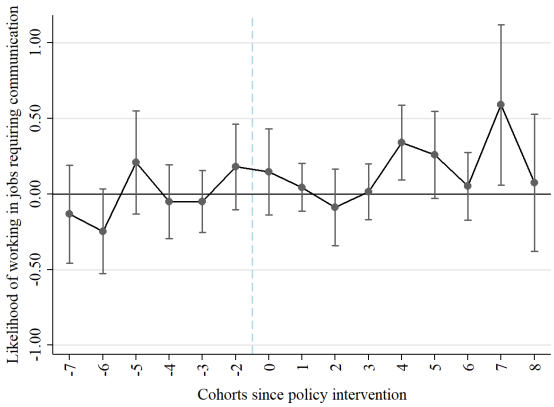
Figure 4: Occupational decisions after exposure to English instruction



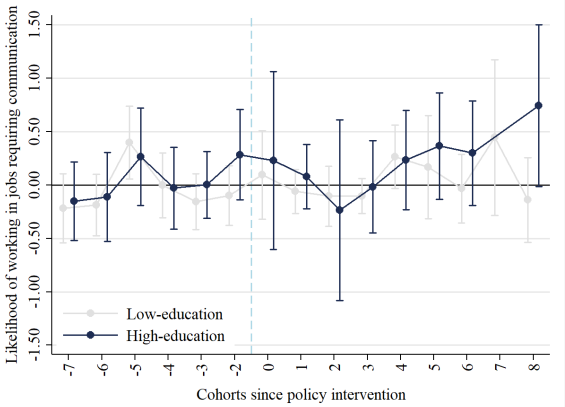
(a) Physically demanding jobs



(b) Physically demanding jobs by education



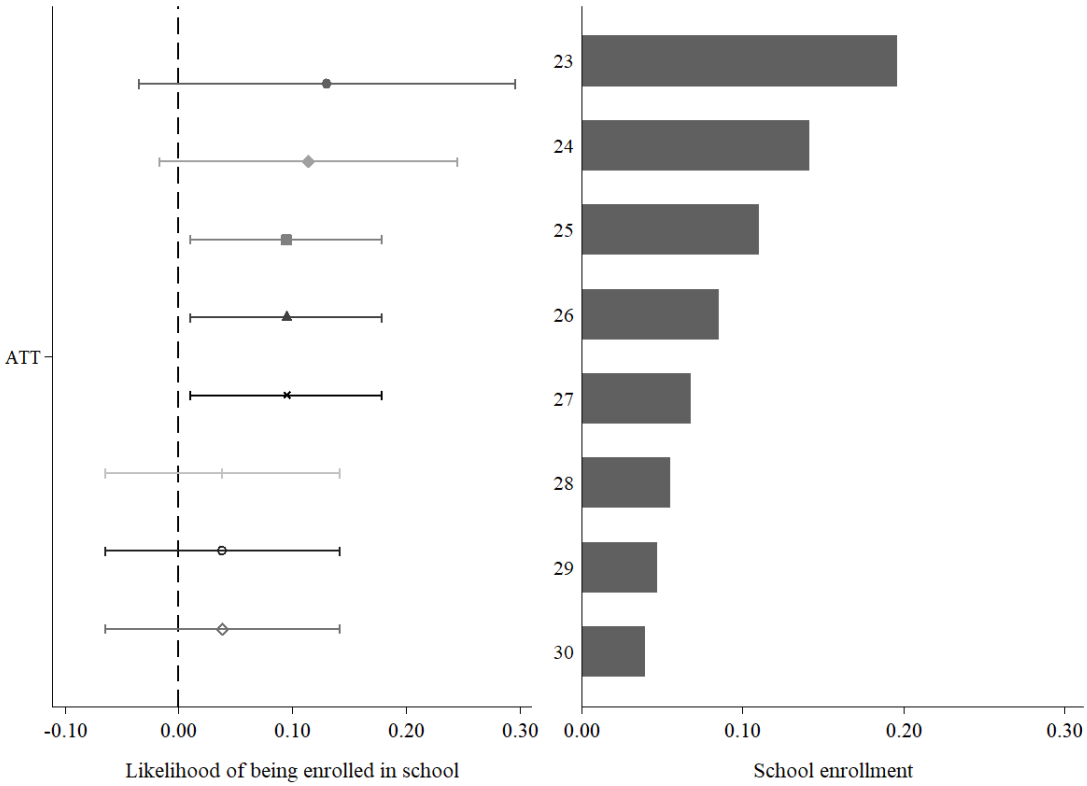
(c) Jobs requiring communication skills



(d) Jobs requiring communication by education

Note: This figure plots robust estimates (as proposed by Callaway and Sant'Anna (2021)) from the interaction terms between the treatment variable and an indicator function for each cohort since the policy intervention, in an event study type regression. The omitted period is -1. The vertical dotted lines indicate the moment of the intervention. The no statistically significant estimates at the left of the vertical dotted line suggest parallel trends before the policy. Confidence intervals at 95% level.

Figure 5: Educational decisions after exposure to English instruction



Note: Each point estimate in the figure on the left corresponds to a different regression, with the difference among regressions being the sample. The initial samples comprise only young cohorts, while subsequent samples include more observations by incorporating older cohorts. It's noteworthy how the impact of the English program diminishes when adding older cohorts, who are less likely to be enrolled in school. However, the effect on enrollment is positive for most specifications utilizing young cohorts.

Table 1: Adult English speaking ability in Mexico

Variable	Full Sample	States w/ English (a)	States wo/ English (b)	Diff. (a-b)
All individuals ages 18-65	6.99	8.65	6.71	1.95***
By gender				
Male	9.46	11.24	9.16	2.08***
Female	5.06	6.58	4.81	1.78***
By age				
18-35	8.27	11.25	7.80	3.45***
36-50	7.01	8.24	6.79	1.44*
51-65	4.46	4.75	4.41	0.34
By educational attainment				
Incomplete primary (0-5 years)	0.82	0.49	0.85	-0.36
Primary school (6 years)	1.48	1.85	1.43	0.42
Lower secondary (7-9 years)	2.42	2.92	2.32	0.60
Upper secondary (10-12 years)	7.45	7.10	7.51	-0.41
College or higher (13-24 years)	21.91	24.94	21.28	3.65**
By ethnicity				
Indigenous	2.46	5.35	2.36	2.98
Non-indigenous	7.28	8.70	7.03	1.67***
By geography				
Urban	8.26	9.55	8.02	1.54***
Rural	2.30	3.28	2.19	1.09**

Note: The sample consist of Mexicans ages 18–65 who self-reported their ability to speak English, using the 2014 Mexican Subjective Well-being Survey (BIARE). The full sample is composed by 33,512 observations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Descriptive statistics

Variable	Full Sample	Speak English (a)	Don't spk English (b)	Diff. (a-b)
Dependent variable				
Wage (monthly pesos)	5,366.88	11,645.27	4,795.18	6,850.09***
Labor supply (hours)	45.97	44.99	46.06	-1.07
Formal job	0.47	0.67	0.45	0.22***
Physically demanding job	0.26	0.10	0.28	-0.18***
Job with comm. skills	0.27	0.58	0.24	0.34***
Independent variables				
English (speaking ability)	0.08	1.00	0.00	-
Hrs English	0.20	0.33	0.18	0.14***
Age (years)	26.81	27.71	26.72	0.99***
Education (years)	10.50	14.16	10.17	4.00***
Female (%)	0.41	0.34	0.41	-0.07**
Indigenous (%)	0.06	0.03	0.06	-0.03***
Married (%)	0.55	0.44	0.57	-0.13***
Rural (%)	0.21	0.09	0.22	-0.13***
Observations	6573	560	6013	6573

Note: These summary statistics consist of Mexicans ages 18–65 who work for pay and self-reported their ability to speak in English. Statistics shown in this table are obtained considering the survey weights. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Returns to English abilities in Mexico

	Full sample					Low-education		High-education	
	(1) ln(wage)	(2) ln(wage)	(3) ln(wage)	(4) ln(wage)	(5) ln(wage)	(6) ln(wage)	(7) ln(wage)	(8) ln(wage)	(9) ln(wage)
Panel A: Men and Women									
Speak Eng	0.807***	0.666***	0.088	-0.082	-0.024	-0.164	-0.556	0.614***	0.075
	(0.147)	(0.138)	(0.144)	(0.139)	(0.144)	(0.515)	(0.533)	(0.131)	(0.164)
Observations	6,573	6,573	6,573	6,573	6,573	3,830	3,830	2,743	2,743
Adjusted R^2	0.011	0.082	0.133	0.181	0.420	0.000	0.524	0.014	0.449
Panel B: Men (β^M)									
Speak Eng	0.561***	0.472**	0.049	-0.028	0.127	-0.858	0.134	0.731***	0.288
	(0.209)	(0.196)	(0.201)	(0.203)	(0.249)	(0.605)	(0.305)	(0.148)	(0.236)
Observations	3,849	3,849	3,849	3,849	3,849	2,401	2,401	1,448	1,448
Adjusted R^2	0.009	0.046	0.091	0.154	0.505	0.008	0.628	0.032	0.538
Panel C: Women (β^W)									
Speak Eng	1.110***	1.014***	0.256	-0.015	-0.112	1.713***	0.840	0.343	-0.379
	(0.219)	(0.209)	(0.223)	(0.224)	(0.294)	(0.276)	(0.741)	(0.236)	(0.329)
Observations	2,724	2,724	2,724	2,724	2,724	1,429	1,429	1,295	1,295
Adjusted R^2	0.012	0.066	0.135	0.213	0.572	0.004	0.688	0.003	0.549
$\beta^M = \beta^W$ [p-value]	[0.089]	[0.059]	[0.524]	[0.946]	[0.649]	[0.000]	[0.125]	[0.159]	[0.152]
Basic controls	NO	YES	YES	YES	YES	NO	YES	NO	YES
Education	NO	NO	YES	YES	YES	NO	YES	NO	YES
Other controls	NO	NO	NO	YES	YES	NO	YES	NO	YES
Locality FE	NO	NO	NO	NO	YES	NO	YES	NO	YES

Note: This table shows the effect of having English abilities on wages in Mexico. The sample consists of Mexicans who were born between 1981 and 1996, and who work for pay. Basic controls include: cohort fixed effects, gender and indigenous dummy. Other controls include: geographical context (rural/urban), marital status and locality fixed effects. Standard errors clustered at locality level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

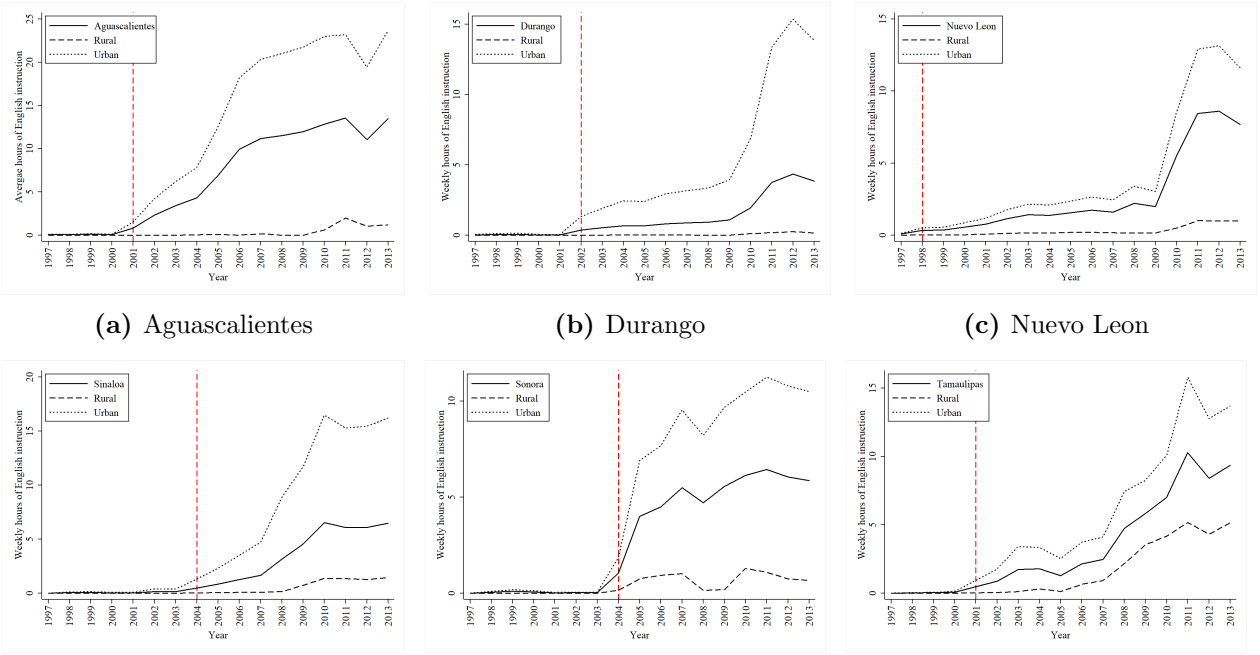
Table 4: Effect of English programs

	(1) Hrs Eng	(2) Speak Eng	(3) ln(wage)	(4) Paid work	(5) Labor supply	(6) Formal work
Panel A: Staggered DiD						
Had Policy	0.546***	0.082*	-0.052	-0.043	-0.046	0.106*
	(0.073)	(0.043)	(0.154)	(0.030)	(0.072)	(0.060)
	[0.000]	[0.034]	[0.727]	[0.144]	[0.526]	[0.089]
Observations	6,573	6,573	6,573	11,965	6,180	6,573
Adjusted R^2	0.681	0.141	0.285	0.258	0.166	0.283
Panel B: Sun and Abraham (2021) interaction weighted estimator						
Had Policy	0.563***	0.092**	-0.120	-0.025	-0.052	0.088*
	(0.058)	(0.024)	(0.133)	(0.025)	(0.066)	(0.052)
Observations	6,264	6,264	6,264	11,813	5,859	6,264
Adjusted R^2	0.666	0.160	0.274	0.257	0.151	0.278
Panel C: Callaway and Sant'Anna (2021)						
Had Policy	0.355***	0.156**	0.769	0.011	-0.051	0.474*
	(0.075)	(0.077)	(0.508)	(0.124)	(0.185)	(0.267)
Observations	6,489	6,489	6,489	10,091	6,110	6,489
Pre-trend test [p-value]	[0.987]	[0.707]	[0.927]	[0.387]	[0.843]	[0.659]
Panel D: Callaway and Sant'Anna (2021): Narrow cohorts, 1985-1995						
Had Policy	0.348***	0.160**	0.774	0.050	-0.049	0.479*
	(0.076)	(0.080)	(0.512)	(0.141)	(0.186)	(0.272)
Observations	4,143	4,143	4,143	7,820	3,889	4,143
Pre-trend test [p-value]	[0.9723]	[0.760]	[0.571]	[0.439]	[0.763]	[0.413]
Mean Dep. Var.	0.103	0.083	7.710	0.541	3.720	0.471

Note: This table shows the effect of state English programs on labor market outcomes. The sample consists of Mexicans who were born between 1981 and 1996. Controls include gender, indigenous people dummy, marital status, education fixed effects, cohort fixed effects, and locality fixed effects. Notice that p-values in brackets are obtained with a wild bootstrap-t procedure with 999 replications and Rademacher weights. On the other hand, in panel C, the null hypothesis of the "pre-trend test" states that all estimates in pre-treatment periods are equal to zero. Standard errors clustered at locality level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix

Figure A.1: Mexican states with English programs



(a) Aguascalientes **(b)** Durango **(c)** Nuevo Leon

(d) Sinaloa **(e)** Sonora **(f)** Tamaulipas

Note: Average weekly hours of English instruction are plotted. The data used comes from the administrative records in the Mexican school census. The vertical dotted lines represent the first year of implementation/expansion of the state's respective English program.

Table A.1: Policy changes in Mexican states

State	Year of impl.	Policy change	Cohorts affected	Hrs of English		Policy details	Comparison state
				Before policy	After policy		
Nuevo Leon	1993	1998	1981-1996	0.97	2.75	Only sixth grades	SLP
Sonora	1993	2004	1989-1996	1.64	5.52	Only 1st and 2nd grades	BC
Tamaulipas	2001	2001	1983-1996	1.21	2.89	Only fourth grades	BC
Aguascalientes	2001	2001	1986-1995	2.36	8.13	No info. available	Zacatecas
Durango	2002	2002	1985-1996	0.33	1.00	Started w/trial stage	SLP
Sinaloa	2004	2004	1989-1996	0.70	1.86	No info. available	Nayarit

Note: These summary statistics consist of Mexicans who where born between 1981–1996 who self-reported their ability to speak English.

Source: I computed the hours of English instruction using the Mexican school census (Statistics 911). Policy details from Nuevo Leon, Sonora, and Tamaulipas were obtained from their respective websites (see subsection 1.4 for the original sources). Details from Durango were obtained from an unofficial source. There are not information available for the state English programs of Aguascalientes and Sinaloa. However, for all states, the information provided from the data in the school census coincides with official and unofficial sources in terms of the release year of each state English program.