

# English skills and labor market outcomes in Mexico

Oscar Gálvez-Soriano

University of Houston  
Department of Economics

April, 2023

# Motivation: Returns to English language abilities

- Language skills are a form of human capital
- English is valuable in the world economy
  - Globalization: trade, technology and information
  - Mobility and better occupations
- In English-speaking countries, English language is needed for communication
- How valuable is English for non-English speaking countries?

# Related Literature

- English speaking countries
  - Immigrants: Bleakley and Chin (2004); Chiswick and Miller (2015)
  - Former British colonies: Azam, Chin and Prakash (2013); Eriksson (2014); Chakraborty and Bakshi (2016)
- Non-English-speaking countries: Lang and Siniver (2009)
- English language skills in Mexico: McConnell and Leclere (2002); Flores-Yeffal (2019); Delgado-Helleseter (2020)

# This paper in a nutshell

## Research Question

- What are the returns to English language skills in a non-English speaking country?

## Strategy

- Describe the prevalence of English skills in Mexico
  - Take advantage of an unusual data set that measures English skills in Mexico
- Quantify the relationship between English skills and labor market outcomes in Mexico
  - Exploit state policy changes that give state-by-cohort variation in exposure to English instruction

# Background

- Importance of English language for Mexico
  - Neighboring country with the US
  - Investment, trade and migration
- Very little is known about English language skills in Mexico
  - Who speaks English?
  - I use the 2014 Subjective Well-being Survey (BIARE) to describe English skills in Mexico
- Very little is known about returns to English skills
  - Public schools spend money on English instruction
  - Some industries and occupations demand workers who speak English

# English speaking ability: a rare skill in Mexico

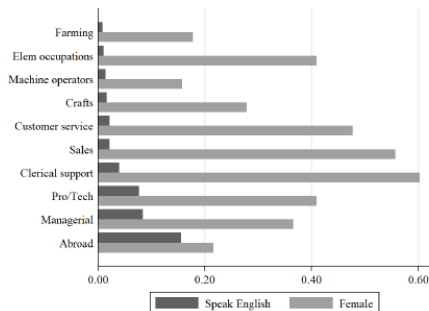
- BIARE is a nationally representative survey with adult respondents 18 and older (ENIGH supplemental survey)
- I use the response to the following question to form a measure of English ability
  - Do you speak English?
  - I code it as one if the respondent says yes, and zero otherwise
- 2.75% of Mexicans speak English



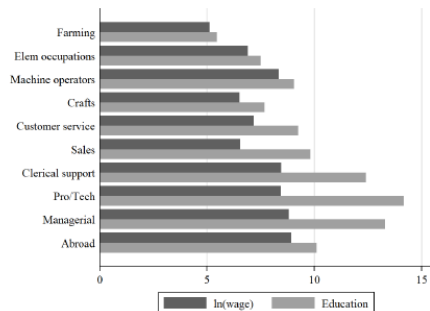
# Where are the English speakers?



# English abilities, wages and education by occupations



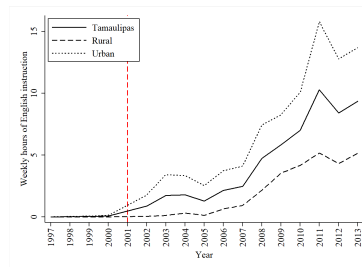
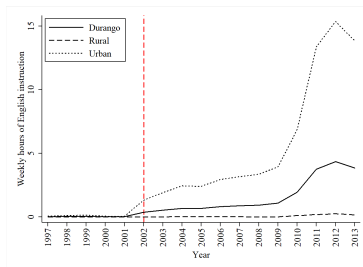
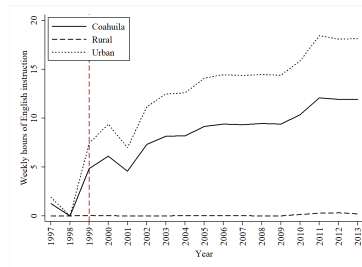
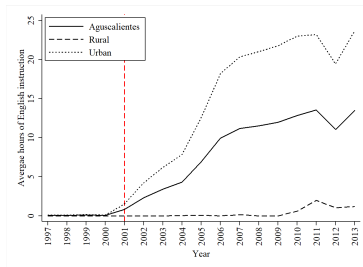
(a) Proportion of female and English speakers



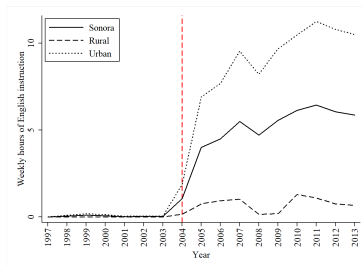
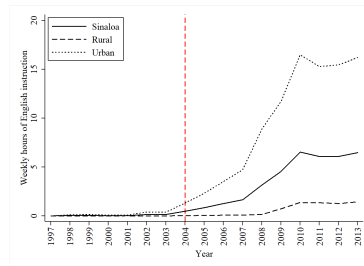
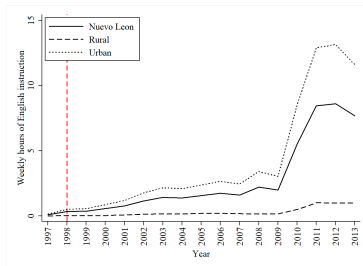
(b) Wages and education



# Policy background



# Policy background



# Empirical framework

We want to estimate the effect of English skills,  $Eng_i$ , on wages,  $\omega_i$ , which can be approximated with the following equation:

$$\omega_i = \alpha + \beta \cdot Eng_i + \mathbf{X}_i \boldsymbol{\Pi} + \epsilon_i$$

where  $\mathbf{X}_i$  is a vector of controls including: education, experience, gender, marital status, ethnicity, student status, cohort FE and geographical context (rural/urban)

# Empirical challenges

- Concern that English skills,  $Eng_i$ , are endogenous in the wage equation
  - Omitted variables: abilities may be correlated with both English skills and wages
  - Measurement error of English skills variable
- OLS estimation would lead to a biased estimate of  $\beta$
- Take advantage of state policy changes in English instruction to form an instrument for English skills to obtain a consistent estimate of  $\beta$
- Since the early 90's some Mexican states implemented English programs to offer English instruction in public primary schools



# Estimating the effect of policy changes

- Use Difference in Differences to estimate the effect of these policies
- Simple illustration with one policy change:

$$y_{isc} = \phi + \gamma \cdot (treatment_s \times after_c) + \delta \cdot treatment_s + \kappa_c + \mathbf{X}_{isc} \mathbf{\Gamma} + \varepsilon_{isc}$$

- where  $y_{isc}$  is the outcome variable; in first stage equation  $y_{isc} = Eng_{isc}$ ; in second stage  $y_{isc} = \omega_{isc}$
- $after_c$ : takes the value of one if the individual  $i$  belongs to one of the cohorts that had exposure to the policy
- $treatment_s$  takes the value of one if individual  $i$  lives in a treated state and zero otherwise

# Parallel Trend Assumption (one policy change)

- Interpreting  $\gamma$  as the effect of the policy requires that the PTA holds
- I offer suggestive evidence on the validity of my identifying assumption using the following event study type specification:

$$y_{isc} = \phi + \sum_c \gamma_c \cdot I_{(treatment_{sc}=c)} + \delta \cdot treatment_s + \kappa_c + \mathbf{X}_{isc} \mathbf{\Gamma} + \varepsilon_{isc}$$

where  $I_{(treatment_{sc}=c)}$  is an indicator function, which identifies if individual  $i$  potentially had exposure, depending on the cohort and state he/she was born. The reference cohort is the one that just missed the policy

# Combining the policy changes (Staggered Difference in Differences)





# Combining the policy changes (Staggered Difference in Differences)

Instead of looking at policies one by one, use all these policies at once

$$y_{isc} = \theta + \psi \cdot HadPolicy_{sc} + \delta_s + \kappa_c + \mathbf{X}_{isc}\Psi + \varepsilon_{isc}$$

where  $HadPolicy_{sc}$  takes the value of one if individual  $i$  lives in a treated state and he/she belongs to one of the affected cohorts (zero otherwise)

# Parallel Trend Assumption (combining all policies)

I use an event study specification to examine if pre-trends are present

$$y_{isc} = \theta + \sum_c \psi_c \cdot I_{(treatment_{sc}=c-c_s^*)} + \delta_s + \kappa_c + \mathbf{X}_{isc}\Psi + \varepsilon_{isc}$$

where  $c_s^*$  denotes the first cohort affected by the intervention in state  $s$ , so  $c - c_s^*$  is the time relative to  $c_s^*$  with negative values reflecting older cohorts not exposed to the policy. The omitted category is -1. Before cohorts with zero effect suggest parallel trends

►► PTA

# IV estimation

Equation of interest (structural equation):

$$\omega_{isc} = \alpha + \beta \cdot Eng_{isc} + \mathbf{X}_{isc}\mathbf{\Pi} + \epsilon_{isc}$$

Use  $HadPolicy_{sc}$  to instrument for  $Eng_{isc}$ . First stage equation:

$$Eng_{isc} = \theta^{fs} + \psi^{fs} \cdot HadPolicy_{sc} + \delta_s^{fs} + \kappa_c^{fs} + \mathbf{X}_{isc}\mathbf{\Psi}^{fs} + \epsilon_{isc}^{fs}$$

Reduced form equation:

$$\omega_{isc} = \theta^{rf} + \psi^{rf} \cdot HadPolicy_{sc} + \delta_s^{rf} + \kappa_c^{rf} + \mathbf{X}_{isc}\mathbf{\Psi}^{rf} + \epsilon_{isc}^{rf}$$

The proposed instrument,  $HadPolicy_{sc}$ , fulfills two conditions:

- Relevance condition
- Exclusion restriction

# Data

## Household survey (2014 BIARE)

- Individual level data
- BIARE surveyed 44,518 households
  - Representative at national and state level
- Very rich questionnaire

## School data on exposure to Eng instruction in primary school

- Mexican School Census (1997-2007)
- Weekly hours of English instruction (exposure)
  - By school-cohort, average over primary school
  - By cohort, take locality average
- Merge English instruction measure to individual-level data (in BIARE) by locality and cohort

# English speakers different from non-Eng speakers

Table : Descriptive statistics

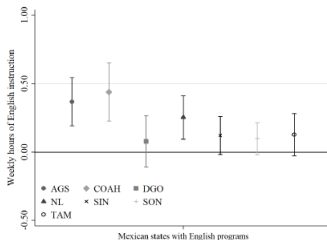
Variable	Full Sample	Speak English (a)	Don't spk English (b)	Diff. (a-b)
<i>Dependent variable</i>				
Wages (monthly pesos)	6,222.40	15,032.45	5,939.54	9,092.91***
<i>Independent variables</i>				
English (speaking ability)	0.03	1.00	0.00	-
Hrs English	0.08	0.12	0.07	0.05***
Age (years)	39.05	38.31	39.08	-0.76**
Education (years)	9.67	14.02	9.53	4.49***
Female (%)	0.38	0.33	0.38	-0.05***
Indigenous (%)	0.06	0.02	0.06	-0.04***
Married (%)	0.66	0.58	0.66	-0.08***
Rural (%)	0.20	0.08	0.20	-0.12***
Observations	49,884	1,664	48,220	49,884

# Results: OLS estimation of structural equation

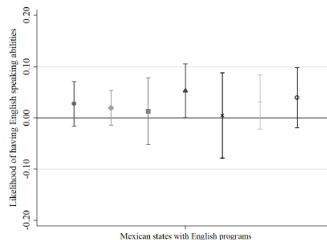
Table : Returns to English abilities in Mexico

	Full sample		Low-education		High-education	
	ln(wage)	ln(wage)	ln(wage)	ln(wage)	ln(wage)	ln(wage)
Speak Eng	0.985*** (0.094)	-0.048 (0.077)	0.004 (0.282)	-0.364 (0.290)	0.586*** (0.079)	0.079 (0.065)
Observations	49,884	49,884	29,043	29,043	20,841	20,841
Adjusted $R^2$	0.005	0.248	0.000	0.253	0.005	0.194
Basic controls	NO	YES	NO	YES	NO	YES
Education	NO	YES	NO	YES	NO	YES
Other controls	NO	YES	NO	YES	NO	YES
Locality FE	NO	YES	NO	YES	NO	YES

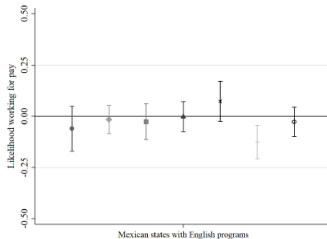
# Results: Effect of state policy changes (DD)



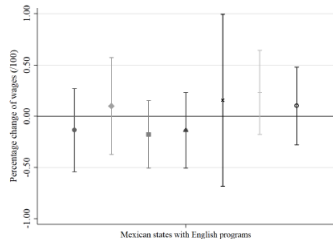
(a) Hours of English



(b) English skills



(c) Paid work



(d) Wages

# Results: Effect of the policies combined (SDD)

Table : Intention to treat effect of offering English instruction at school  
(SDD estimate)

	(1)	(2)	(3)	(4)	(5)
	Hrs Eng	Speak Eng	ln(wage)	Paid work	Student
<i>Panel A: Full sample</i>					
Had Policy	0.308*** (0.046)	0.015** (0.008)	0.019 (0.080)	-0.020 (0.013)	0.038*** (0.014)
Observations	13,131	13,131	13,131	22,517	22,517
Adjusted $R^2$	0.596	0.075	0.153	0.219	0.370



# Results: Effect of the policies combined (SDD)

Table : Intention to treat effect of offering English instruction at school  
(SDD estimate)

	(1)	(2)	(3)	(4)	(5)
	Hrs Eng	Speak Eng	ln(wage)	Paid work	Student
<i>Panel A: Full sample</i>					
Had Policy	0.308*** (0.046)	0.015** (0.008)	0.019 (0.080)	-0.020 (0.013)	0.038*** (0.014)
Observations	13,131	13,131	13,131	22,517	22,517
Adjusted $R^2$	0.596	0.075	0.153	0.219	0.370

# Results: Effect of the policies combined (SDD)

Table : Intention to treat effect of offering English instruction at school  
(SDD estimate)

	(1)	(2)	(3)	(4)	(5)
	Hrs Eng	Speak Eng	ln(wage)	Paid work	Student
<i>Panel A: Full sample</i>					
Had Policy	0.308*** (0.046)	0.015** (0.008)	0.019 (0.080)	-0.020 (0.013)	0.038*** (0.014)
Observations	13,131	13,131	13,131	22,517	22,517
Adjusted $R^2$	0.596	0.075	0.153	0.219	0.370

# Results: Effect of the policies combined (SDD)

Table : Intention to treat effect of offering English instruction at school  
(SDD estimate)

	(1)	(2)	(3)	(4)	(5)
	Hrs Eng	Speak Eng	ln(wage)	Paid work	Student
<i>Panel A: Full sample</i>					
Had Policy	0.308*** (0.046)	0.015** (0.008)	0.019 (0.080)	-0.020 (0.013)	0.038*** (0.014)
Observations	13,131	13,131	13,131	22,517	22,517
Adjusted $R^2$	0.596	0.075	0.153	0.219	0.370

# Results: Gender heterogeneous effects (SDD)

Table : Intention to treat effect of offering English instruction at school (SDD estimate)

	(1) Hrs Eng	(2) Speak Eng	(3) ln(wage)	(4) Paid work	(5) Student
<i>Panel B: Men (<math>\beta^M</math>)</i>					
Had Policy	0.295*** (0.046)	0.016 (0.011)	-0.044 (0.069)	-0.023 (0.018)	0.044*** (0.017)
Observations	8,008	8,008	8,008	11,021	11,021
Adjusted $R^2$	0.581	0.069	0.186	0.313	0.408
<i>Panel C: Women (<math>\beta^W</math>)</i>					
Had Policy	0.331*** (0.057)	0.014 (0.014)	0.152 (0.152)	-0.003 (0.023)	0.031* (0.017)
Observations	5,123	5,123	5,123	11,496	11,496
Adjusted $R^2$	0.604	0.056	0.150	0.158	0.326
$\beta^M = \beta^W$ [p-value]	[0.208]	[0.870]	[0.182]	[0.443]	[0.445]

### Results: Educational heterogeneous effects (SDD)

Table 1: Intention to treat effect of offering English instruction at school  
(SDD estimate)

	(1)	(2)	(3)	(4)	(5)
	Hrs	Speak	ln(wage)	Paid	Student
	Eng	Eng		work	
<i>Panel D: Low education sample (<math>\beta^L</math>)</i>					
Had Policy	0.305***	0.010	-0.169	0.016	0.005
	(0.062)	(0.007)	(0.123)	(0.021)	(0.014)
Observations	6,624	6,624	6,624	10,898	10,898
Adjusted $R^2$	0.522	0.016	0.162	0.225	0.043
<i>Panel E: High education sample (<math>\beta^H</math>)</i>					
Had Policy	0.304***	0.019	0.227**	-0.025	0.031*
	(0.040)	(0.014)	(0.099)	(0.018)	(0.017)
Observations	6,507	6,507	6,507	11,619	11,619
Adjusted $R^2$	0.659	0.070	0.146	0.263	0.432
$\beta^L = \beta^H$ [p-value]	[0.973]	[0.590]	[0.005]	[0.119]	[0.224]

## ►► Geographical context

# Results: IV estimate on wages

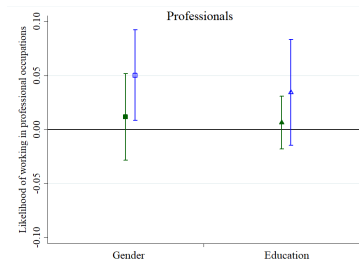
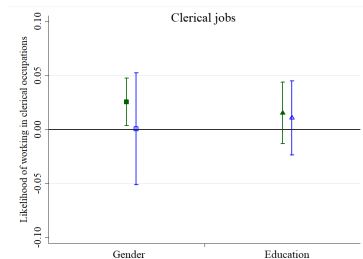
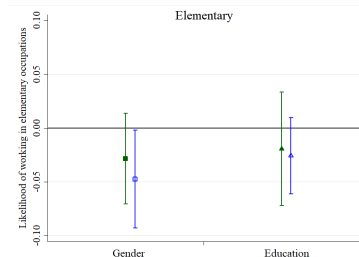
Table : Returns to English abilities  
(IV estimate)

	(1) Structural-OLS	(2) First Stage	(3) Reduced Form	(4) Structural-IV
Speak Eng	0.142 (0.102)			1.252 (5.084)
Had Policy		0.015** (0.008)	0.019 (0.080)	
Observations	13,131	13,131	13,131	13,131
Adjusted $R^2$	0.153	0.075	0.153	

# Mechanisms

- Cognitive skills
  - Acquisition of English skills
  - No effect on other skills: Language and Mathematics  
(Galvez-Soriano, 2023)
- Occupational choices
  - Move into occupations that require English skills
  - Better working conditions

# Mechanisms: Occupational choices (SDD) ▶ All





# Robustness checks

- Concern about SDD estimator in the presence of heterogeneous treatment effects [» SDD](#)
- Different comparison group (all neighboring states) [» DD](#)
- Narrower cohorts [» SDD](#)

# Discussion

- No returns to English skills in Mexico
  - Young adults (18-24) are still enrolled in school
  - Potentially high-earners
- However, exposure to English instruction...
  - Leads to the acquisition of English abilities
  - Increases school enrollment
  - Improves working conditions

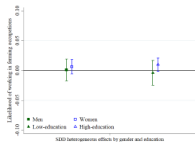
# Results: Ethnicity heterogeneous effects

Table : Heterogeneous effects of offering English instruction at school  
(SDD estimate)

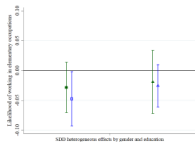
	(1)	(2)	(3)	(4)	(5)
	Hrs Eng	Speak Eng	ln(wage)	Paid work	Student
<i>Panel A: By ethnicity</i>					
<b>Indigenous (<math>\beta^I</math>)</b>					
Had Policy	0.110 (0.160)	-0.034 (0.091)	0.309 (1.202)	0.076 (0.187)	0.286** (0.126)
Observations	188	188	188	340	340
Adjusted $R^2$	0.929	0.410	0.193	0.380	0.329
<b>Non-Indigenous (<math>\beta^N</math>)</b>					
Had Policy	0.306*** (0.046)	0.015* (0.008)	0.030 (0.081)	-0.020 (0.013)	0.036*** (0.014)
Observations	12,943	12,943	12,943	22,177	22,177
Adjusted $R^2$	0.593	0.076	0.148	0.219	0.371
$\beta^I = \beta^N$ [p-value]	[0.348]	[0.208]	[0.212]	[0.819]	[0.011]



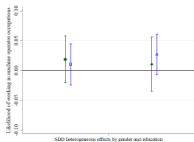
# Mechanisms: Occupational choices (SDD) ▶▶ Back



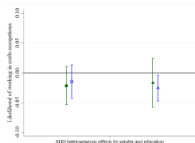
(a) Farming



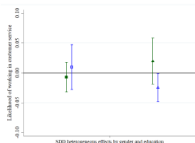
(b) Elementary



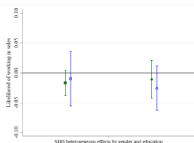
(c) Machine operator



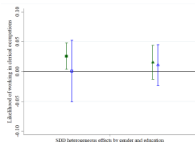
(d) Crafts



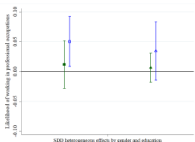
(e) Customer service



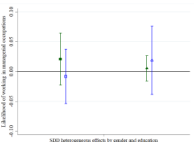
(f) Sales



(g) Clerks



(h) Professionals



(i) Managerial

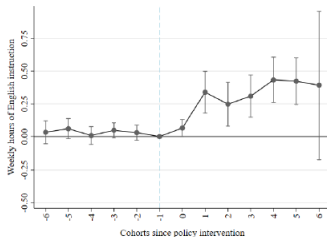
# SDD correction: Sun, Liyang and Sarah Abraham (2021)

[▶▶ Back](#)

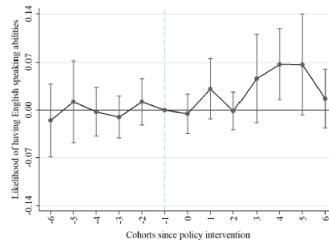
**Table :** Intention to Treat effect of offering English instruction at school (SDD estimate)

	(1) Hrs Eng	(2) Speak Eng	(3) ln(wage)	(4) Paid work	(5) Student
Had Policy	0.323*** (0.045)	0.015** (0.007)	0.021 (0.081)	-0.015 (0.012)	0.035*** (0.013)
Observations	13,063	13,063	13,063	22,493	22,493
Adjusted $R^2$	0.616	0.073	0.148	0.219	0.371

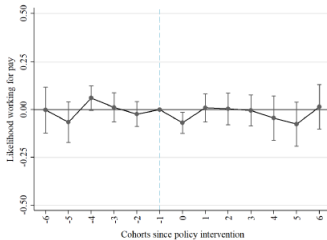
# Robustness check: Narrower cohort window

[▶▶ Back](#)


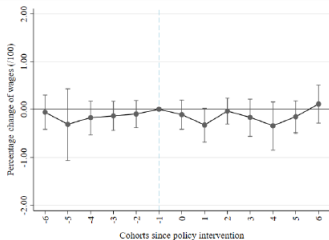
(a) Hours of English



(b) Speak English



(c) Paid work



(d) Ln(wage)

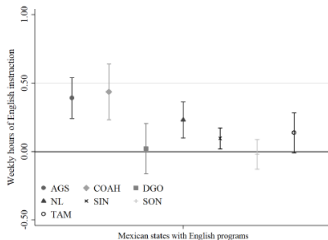
# Robustness check: Narrower cohort window [» Back](#)

Table : Returns to English abilities  
(IV estimate with narrower comparison group)

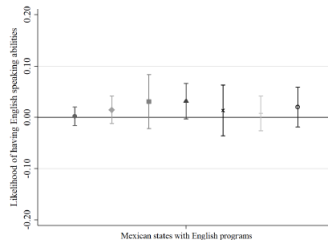
	(1) Structural-OLS	(2) First Stage	(3) Reduced Form	(4) Structural-IV
Speak Eng	0.086 (0.156)			-4.081 (6.348)
Had Policy		0.018* (0.010)	-0.073 (0.111)	
Observations	5,926	5,926	5,926	5,926
Adjusted $R^2$	0.159	0.050	0.159	



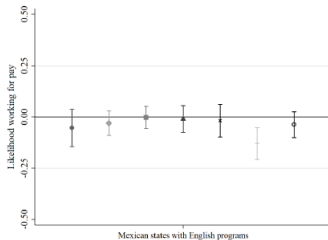
# Robustness check: Different control groups

[▶ Back](#)

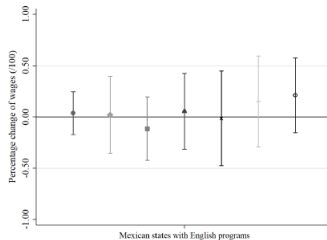
(a) Hours of English



(b) English skills

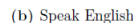
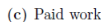
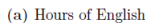


(c) Paid work

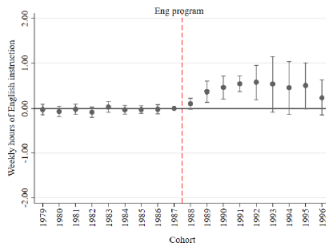


(d) Wages

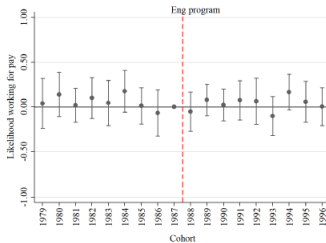
►► Back



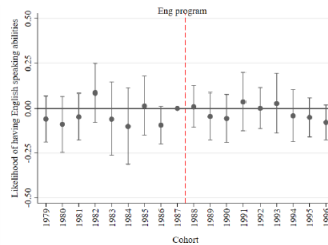
# Parallel Trend Assumption (Coahuila)



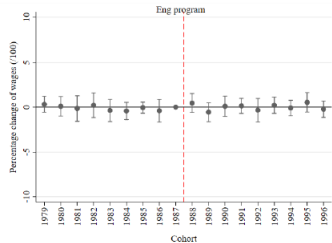
(a) Hours of English



(c) Paid work

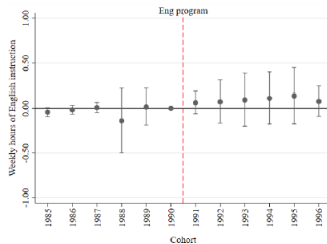


(b) Speak English

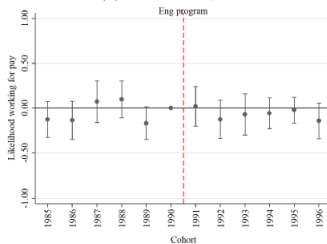


(d) Ln(wage)

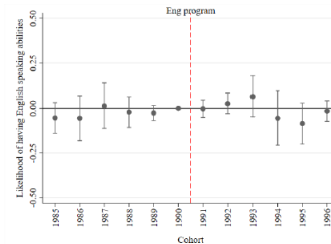
# Parallel Trend Assumption (Durango)



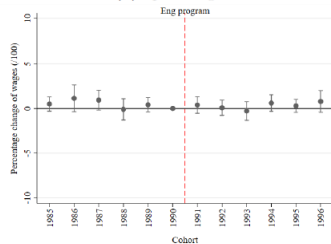
(a) Hours of English



(b) Speak English

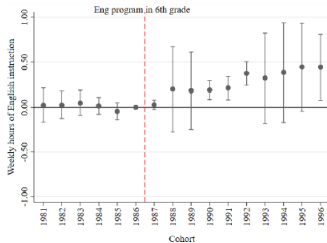


(c) Paid work

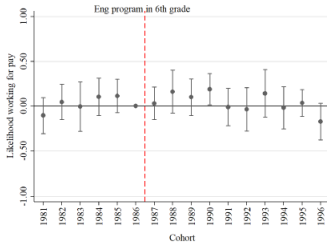


(d) Ln(wage)

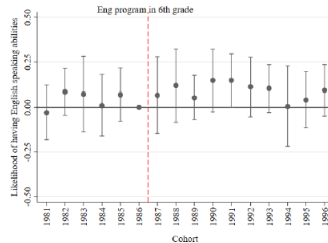
# Parallel Trend Assumption (Nuevo Leon)



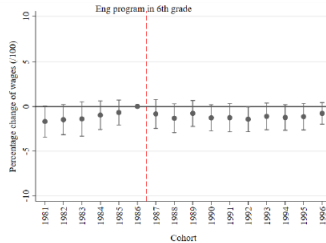
(a) Hours of English



(c) Paid work

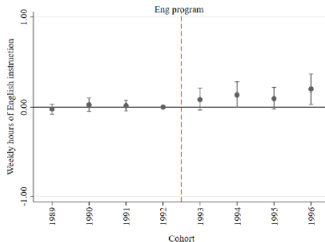


(b) Speak English

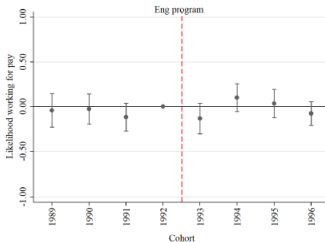


(d) Ln(wage)

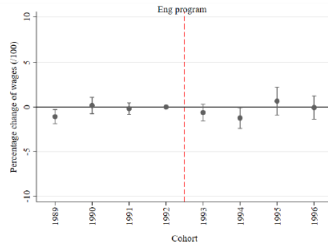
# Parallel Trend Assumption (Sinaloa)



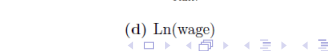
(a) Hours of English



(b) Speak English

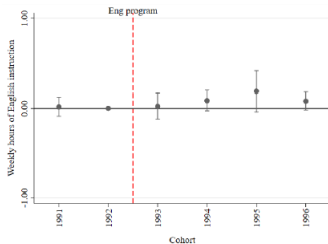


(c) Paid work

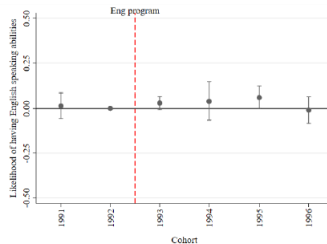


(d) Ln(wage)

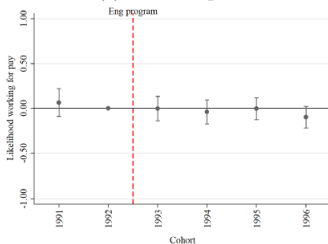
# Parallel Trend Assumption (Sonora)



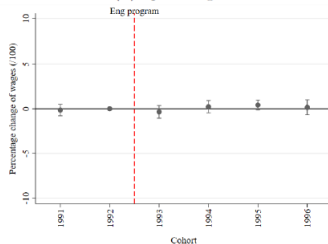
(a) Hours of English



(b) Speak English

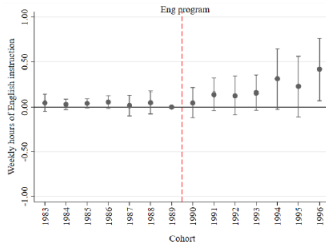


(c) Paid work

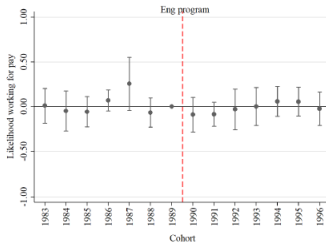


(d) Ln(wage)

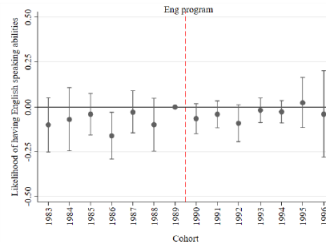
# Parallel Trend Assumption (Tamaulipas)

[▶▶ Back](#)

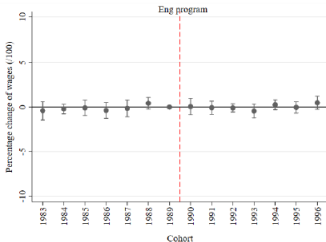
(a) Hours of English



(c) Paid work



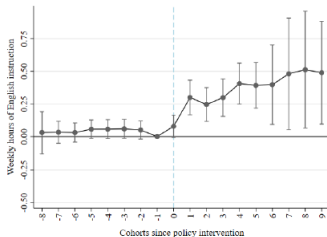
(b) Speak English



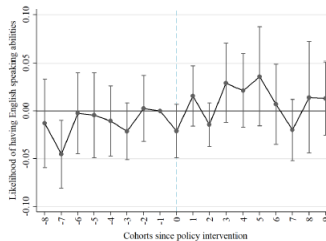
(d) Ln(wage)



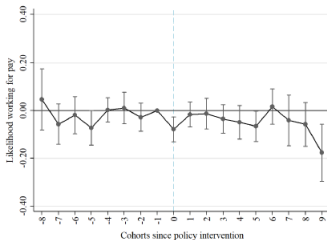
# PTA Staggered DiD: All states

[▶▶ Back](#)


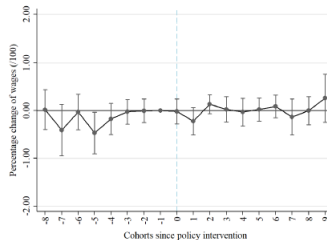
(a) Hours of English



(b) Speak English



(c) Paid work



(d) Ln(wage)