

# English skills and labor market outcomes in Mexico

Oscar Gálvez-Soriano

University of Chicago  
Department of Economics

September, 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

## 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)

# This paper in a nutshell

## Research question

- What are the returns to English language skills in Mexico?

## What I do

- 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
- Very little is known about returns to English skills in Mexico
  - I use the 2014 Subjective Well-being Survey (BIARE)

# 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
- 7% of Mexicans speak English



# Empirical framework

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

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

where  $\mathbf{X}_i$  is a vector of controls including: education, experience, gender, marital status, ethnicity, cohort FE and locality FE

# 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  $\beta$

# States with and without the policy





# Staggered Difference in Differences

I examine all these policies at once, using the following specification:

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

where  $\text{HadPolicy}_{sc}$  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 (SDD)

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

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

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.  $I_{(k=c-c_s^*)}$  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

►► PTA

# IV estimation

Equation of interest:

$$\omega_{isc} = \alpha + \beta Eng_{isc} + \delta_s + \kappa_c + \mathbf{X}_{isc}\Psi + \varepsilon_{isc}$$

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

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

Reduced form equation:

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

# Data

## » Descriptive Stats

### Household survey (2014 BIARE)

- Individual level data (18-38 years old)
- BIARE surveyed 44,518 households
  - Representative at national and state level
- Very rich questionnaire, including English skills

### School data on exposure to English 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

# Results: Effect of English policies

**Table 4:** Effect of English programs

	(1)	(2)	(3)	(4)
	Hrs Eng	Speak Eng	ln(wage)	Paid work
<i>Panel A: Staggered DiD</i>				
Had Policy	0.331*** (0.058)	0.028* (0.017)	-0.093 (0.129)	-0.002 (0.023)
Observations	5,437	5,437	5,437	8,979
Adjusted $R^2$	0.573	0.177	0.172	0.230
Mean Dep. Var.	0.119	0.106	7.972	0.606

# Results: Effect of English policies

**Table 4:** Effect of English programs

	(1)	(2)	(3)	(4)
	Hrs Eng	Speak Eng	ln(wage)	Paid work
<i>Panel A: Staggered DiD</i>				
Had Policy	0.331*** (0.058)	0.028* (0.017)	-0.093 (0.129)	-0.002 (0.023)
Observations	5,437	5,437	5,437	8,979
Adjusted $R^2$	0.573	0.177	0.172	0.230
Mean Dep. Var.	0.119	0.106	7.972	0.606

# Results: Effect of English policies

**Table 4:** Effect of English programs

	(1)	(2)	(3)	(4)
	Hrs Eng	Speak Eng	ln(wage)	Paid work
<i>Panel A: Staggered DiD</i>				
Had Policy	0.331*** (0.058)	0.028* (0.017)	-0.093 (0.129)	-0.002 (0.023)
Observations	5,437	5,437	5,437	8,979
Adjusted $R^2$	0.573	0.177	0.172	0.230
Mean Dep. Var.	0.119	0.106	7.972	0.606

# Results: IV estimate of returns to English abilities

**Table 5:** Returns to English abilities  
(IV estimate)

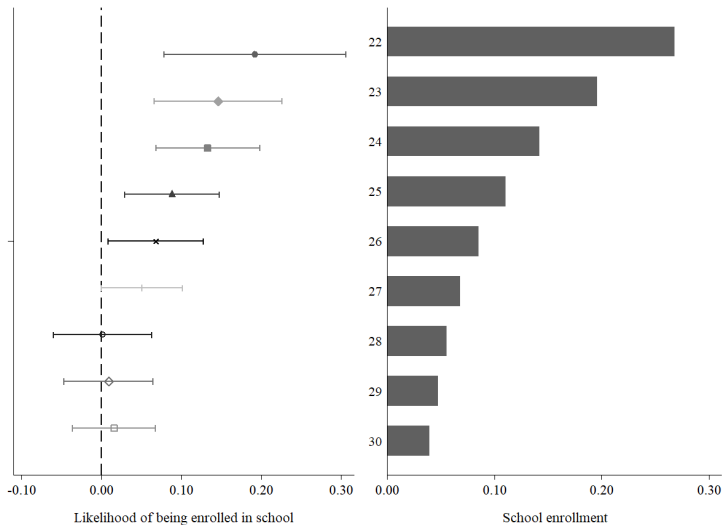
	(1) Structural-OLS	(2) First Stage	(3) Reduced Form	(4) Structural-IV
Speak Eng	0.061 (0.110)			-3.285 (4.548)
Had Policy		0.028* (0.017)	-0.093 (0.129)	
Observations	5,437	5,437	5,437	5,437
Adjusted $R^2$	0.172	0.177	0.172	



# Mechanisms

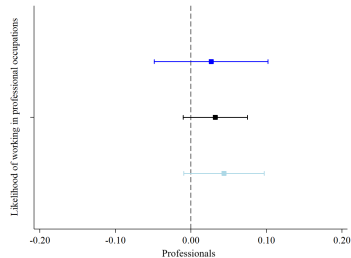
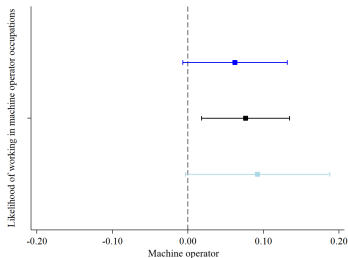
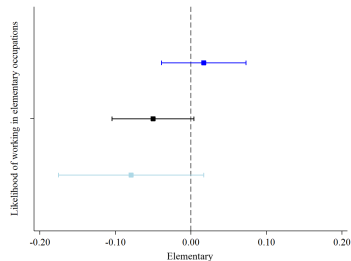
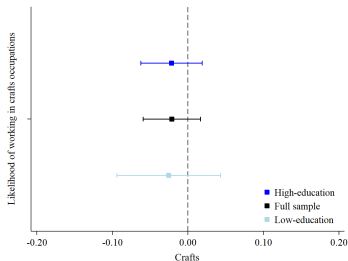
- Cognitive skills
  - Acquisition of English skills
  - No effect on other skills: Language and Mathematics  
(Gálvez-Soriano, 2023)
- School enrollment
  - Negative effect on wages in the short-run, but positive in the long-run?
- Occupational choices
  - Move into occupations that require English skills
  - Better working conditions

# Mechanisms: School enrollment

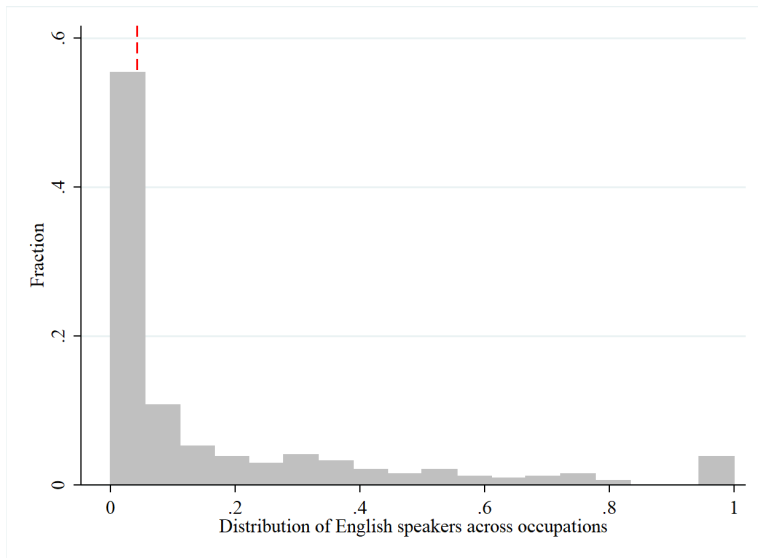


# Mechanisms: Occupational choices

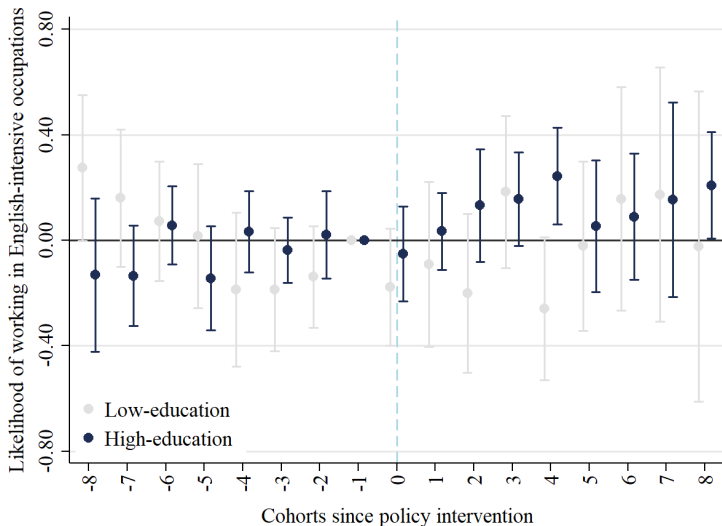
► All



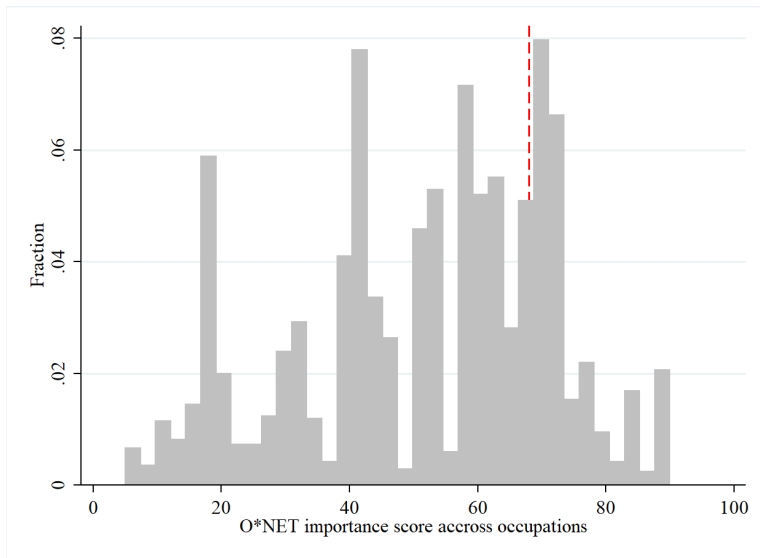
# More likely to work in English-intensive jobs?



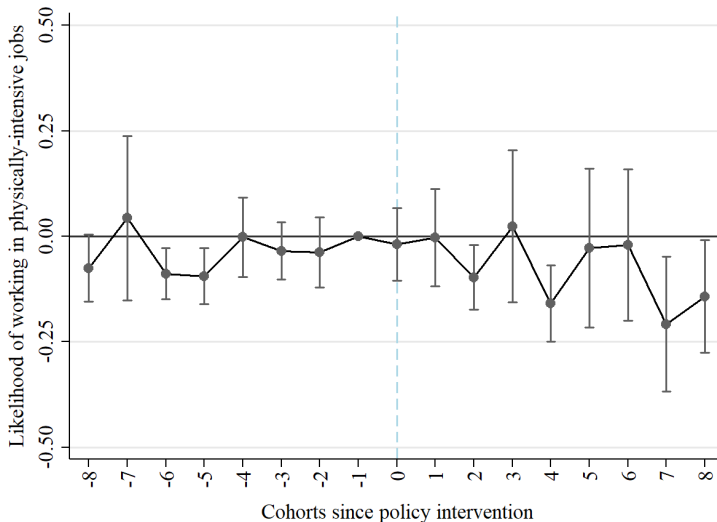
# More likely to work in English-intensive jobs?



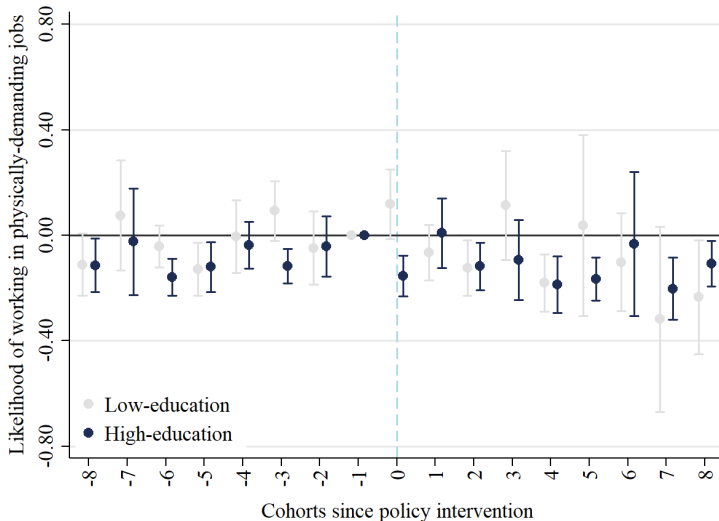
# Less likely to work in physically demanding jobs?



# Less likely to work in physically demanding jobs?

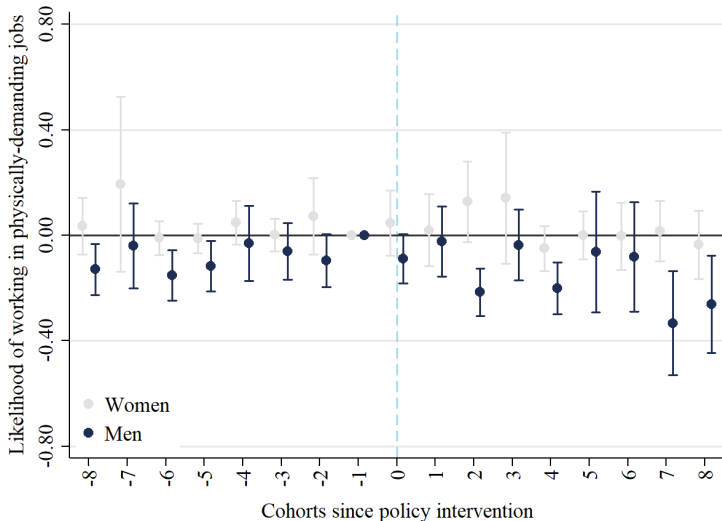


# Less likely to work in physically demanding jobs?





# Less likely to work in physically demanding jobs?



# Robustness checks

- Concern about staggered DiD estimator in the presence of heterogeneous treatment effects ▶ SDD
  - Sun and Abraham (2021)
  - Callaway and Sant'Anna (2021)
- Narrower cohorts ▶ SDD

# Conclusion

- First study to examine English skills and labor market outcomes in Mexico using large nationally representative sample
- I use variation in English skills generated by state policy changes
- I find no effect on wages, shifts across occupations. Highly educated are:
  - more likely to work in English intensive jobs
  - less likely to work in physically demanding jobs

# Thank you!

*For more about me and my research, please scan here:*

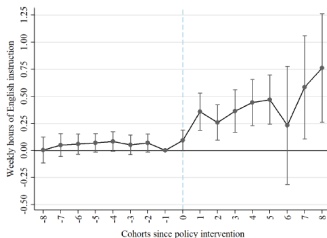


# English speakers different from non-Eng speakers [▶ Back](#)

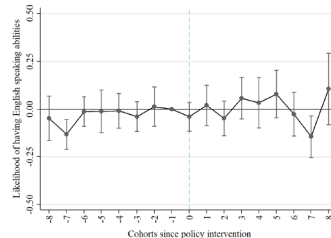
Table 2: Descriptive statistics

Variable	Full Sample	Speak English (a)	Don't spk English (b)	Diff. (a-b)
<i>Dependent variable</i>				
Wage (monthly pesos)	6,261.81	15,042.61	5,529.11	9,513.50***
Farming	0.07	0.02	0.08	-0.06***
Elementary	0.23	0.08	0.24	-0.16***
Machine operators	0.09	0.04	0.10	-0.05***
Crafts	0.10	0.05	0.11	-0.06***
Customer service	0.08	0.05	0.08	-0.03***
Sales	0.12	0.08	0.12	-0.04***
Clerical support	0.06	0.08	0.06	0.02**
Professionals	0.16	0.37	0.14	0.23***
Managerial	0.08	0.21	0.07	0.14***
Abroad	0.00	0.02	0.00	0.01***
<i>Independent variables</i>				
English (speaking ability)	0.08	1.00	0.00	-
Hrs English	0.07	0.12	0.06	0.05***
Age (years)	39.54	38.35	39.64	-1.29***
Education (years)	9.68	14.02	9.31	4.71***
Female (%)	0.41	0.33	0.42	-0.09***
Indigenous (%)	0.06	0.02	0.07	-0.04***
Married (%)	0.63	0.58	0.63	-0.06***
Rural (%)	0.20	0.08	0.21	-0.13***
Observations	20,492	1,658	18,834	20,492

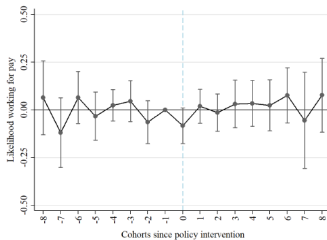
# PTA Staggered DiD: All states

[▶ Back](#)

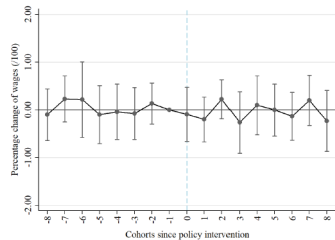
(a) Hours of English



(b) Speak English



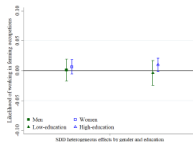
(c) Paid work



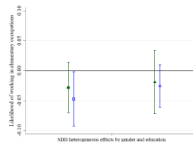
(d) Ln(wage)

# Mechanisms: Occupational choices

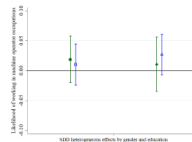
► Back



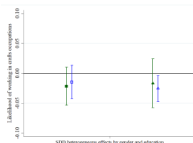
(a) Farming



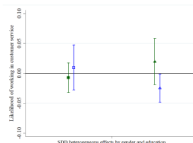
(b) Elementary



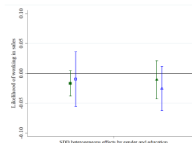
(c) Machine operator



(d) Crafts



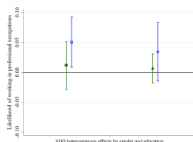
(e) Customer service



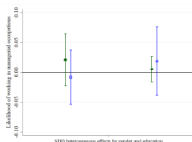
(f) Sales



(g) Clerks



(h) Professionals



(i) Managerial

# Staggered DiD correction

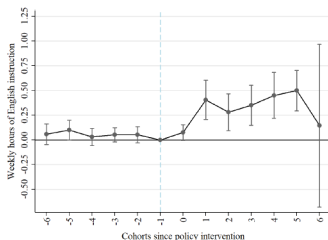
[» Back](#)

Table 4: Effect of English programs

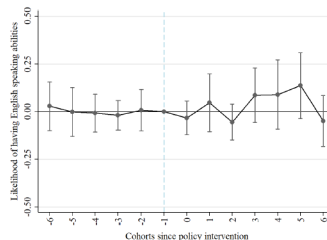
	(1)	(2)	(3)	(4)
	Hrs	Speak	ln(wage)	Paid
	Eng	Eng		work
<i>Panel B: Sun and Abraham (2021) interaction weighted estimator</i>				
Had Policy	0.336***	0.028*	-0.096	0.000
	(0.055)	(0.015)	(0.123)	(0.021)
Observations	5,339	5,339	5,339	8,917
Adjusted $R^2$	0.600	0.176	0.159	0.229
<i>Panel C: Callaway and Sant'Anna (2021)</i>				
Had Policy	0.327***	0.068	0.001	0.072
	(0.070)	(0.044)	(0.195)	(0.050)
Observations	5,418	5,418	5,418	8,979
Pre-trend Test [p-value]	[0.000]	[0.1719]	[0.2747]	[0.2006]
Mean Dep. Var.	0.119	0.106	7.972	0.606



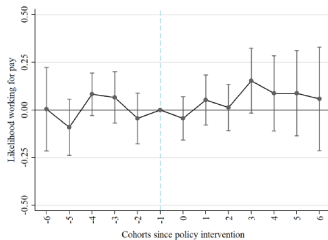
# Robustness check: Narrower cohort window

[▶ Back](#)

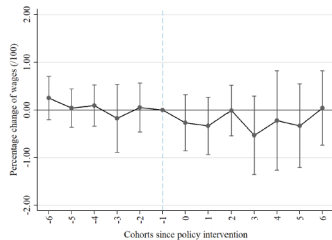
(a) Hours of English



(b) Speak English



(c) Paid work



(d) Ln(wage)

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

**Table 6:** Returns to English abilities  
(IV estimate with narrower comparison group)

	(1)	(2)	(3)	(4)
	Structural-OLS	First Stage	Reduced Form	Structural-IV
Speak Eng	0.011 (0.190)			-11.824 (14.160)
Had Policy		0.022 (0.029)	-0.257 (0.189)	
Observations	2,283	2,283	2,283	2,283
Adjusted $R^2$	0.171	0.147	0.173	