Seeing is Believing: Inequality and the Impact of Television on Hispanics

Andrew Kao
Harvard

March 2021

Motivation

- Hispanics in the US face large obstacles. Compared to white people, they are:
  - **60%** less likely to complete college,
  - **68%** less likely to found a business,
  - and earn **24%** less than white men

Motivation

- Hispanics in the US face large obstacles. Compared to white people, they are:
  - **60%** less likely to complete college,
  - **68%** less likely to found a business,
  - and earn **24%** less than white men
- **50%** of Hispanics watch satellite or broadcast Spanish Language TV (SLTV)

Motivation

- Hispanics in the US face large obstacles. Compared to white people, they are:
  - **60%** less likely to complete college,
  - **68%** less likely to found a business,
  - and earn **24%** less than white men
- **50%** of Hispanics watch satellite or broadcast Spanish Language TV (SLTV)
- Large literature on how TV affects behavior (Yanigazawa-Drott 2014; DellaVigna & al. 2007; Ferrara & al., 2012)
- Prior efforts to study Hispanic interaction with media focused on politics (Waldfogel & al. 2009; Trujillo & al. 2012)

This project:

Examine the effect of SLTV in two domains:

1. Education: Hispanic performance in schools
2. Commercial: Hispanic firm ownership

- Goal: To show that SLTV improves outcomes for Hispanics, and that these effects (at least partially) stem from an identity mechanism

This project:

Examine the effect of SLTV in two domains:

1. Education: Hispanic performance in schools
2. Commercial: Hispanic firm ownership

- Goal: To show that SLTV improves outcomes for Hispanics, and that these effects (at least partially) stem from an identity mechanism
- Identification: follow Velez & Newman (2019) and construct spatial RD arising from FCC TV signal regulation

Contribution

- Existing work on Hispanic communities often geographically constrained & media studies only concerned with effect on politics (Velez & Newman 2019; Trujillo & al. 2012).

→ Identify causal effect on larger scale and with more granularity (geocoded microdata)

→ Provide a first look at how media affects business and schooling outcomes for Hispanics

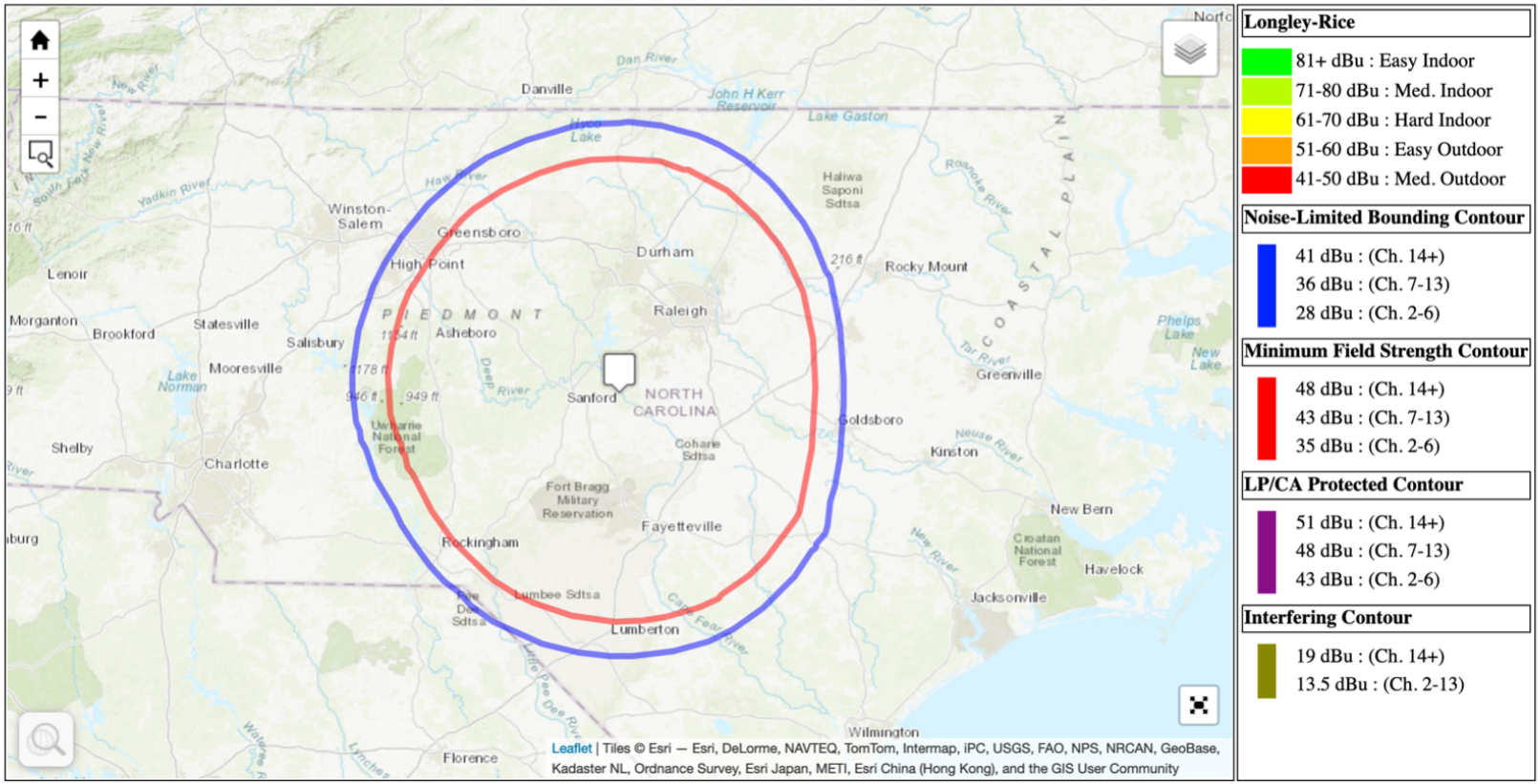
- Existing research that shows identity is a powerful mechanism driving meaningful outcomes (Benjamin & al. 2007; Bursztyn & al. 2015). New research on how identity is constructed and strengthened (Atkin & al. 2019; Bazzi & al. 2019)

→ Show how identity can be bolstered by the media and how it can help reduce inequality

Coverage Map for TV Station WUVC-DT

Coverage Maps

WUVC-DT (40-1) BLCDT-20060912ACZ



Longley-Rice

- 81+ dBu : Easy Indoor
- 71-80 dBu : Med. Indoor
- 61-70 dBu : Hard Indoor
- 51-60 dBu : Easy Outdoor
- 41-50 dBu : Med. Outdoor

Noise-Limited Bounding Contour

- 41 dBu : (Ch. 14+)
- 36 dBu : (Ch. 7-13)
- 28 dBu : (Ch. 2-6)

Minimum Field Strength Contour

- 48 dBu : (Ch. 14+)
- 43 dBu : (Ch. 7-13)
- 35 dBu : (Ch. 2-6)

LP/CA Protected Contour

- 51 dBu : (Ch. 14+)
- 48 dBu : (Ch. 7-13)
- 43 dBu : (Ch. 2-6)

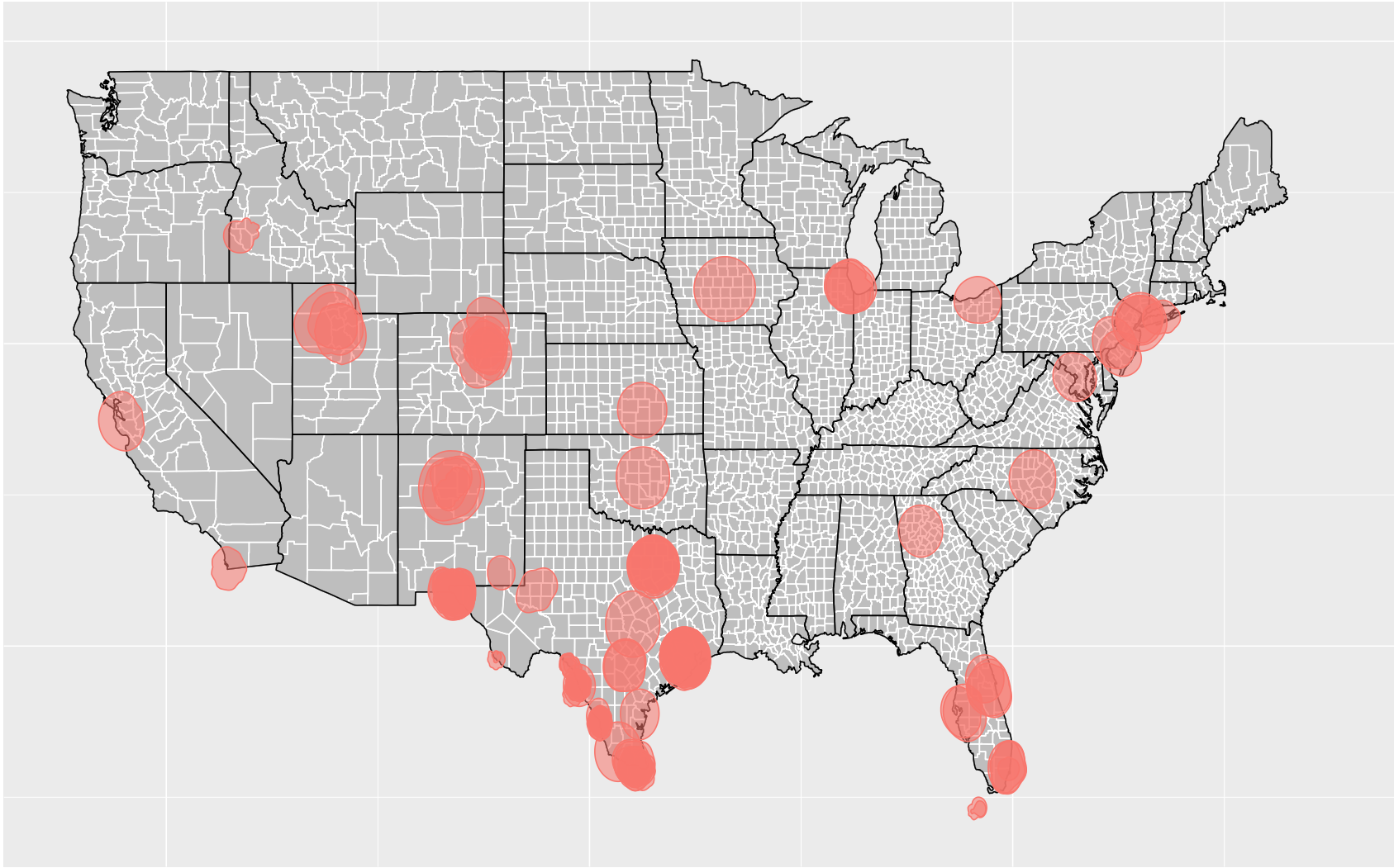
Interfering Contour

- 19 dBu : (Ch. 14+)
- 13.5 dBu : (Ch. 2-13)

Empirical Strategy

- OET Bulletin No. 69 — protect TV stations in (50,90) coverage contour areas
  - Mechanical formula based on geographic/technical factors (not political/economic)
  - Fairly large boundaries that typically cut through small towns/suburbs
  - Purchase/constructed antennas prior to 1977
- Spanish Language TV: Isolate effect on Hispanic communities
- Both RD and instrument with distance
  - Keep observations within 100 KM of boundary for comparability
  - Fuzzy RD: only protection from interference + distance effect

SLTV Coverage Maps for the USA



Empirical Specification

Main Model:

$$Y_i = \beta_0 + \beta \mathbb{I}[InsideContour_i] \times Distance_i + \gamma X_i + \epsilon_i$$

where i is a unit of observation (i.e. a school) or one tile on a grid map (with firms)

Place focus on the RD, dummy for whether observation falls inside contour

Data - General

- Instrument:
  - Identify 100 Spanish Language TV stations across the US from TMS
  - Station contours and other station data from the FCC (use data from 2015 for consistency with outcomes)
- Geocoding:
  - ArcGIS: 99%+ successfully geocoded, but data limit (schools)
  - US Census Geocoder: 80% successfully geocoded (firms)
- Demographic and migration information at county level from ACS

Selection? Migration?

	IHS(# Hispanic Migrants)		
Panel A: Origin County Inside Contour	(1)	(2)	(3)
Dummy: Destination Outside TV Contour	-0.387^{***}	-0.286^{***}	-0.280^{***}
	(0.048)	(0.044)	(0.044)
TV Dummy $\times$ Distance to Origin	-0.003^{**}	-0.004^{***}	-0.004^{***}
	(0.001)	(0.001)	(0.001)
TV Dummy $\times$ Distance to Destination	0.001	-0.002^{*}	-0.002
	(0.001)	(0.001)	(0.001)
Distance from Contour to Origin (KM)	0.001	0.003^{*}	0.003
	(0.002)	(0.002)	(0.002)
Distance from Contour to Destination (KM)	-0.001	0.002	0.002
	(0.001)	(0.001)	(0.001)
Observations	8,479	8,479	8,479
Panel B: Origin County Outside Contour			
Dummy: Destination Inside TV Contour	-0.078	-0.123	-0.120
	(0.108)	(0.096)	(0.096)
TV Dummy $\times$ Distance to Origin	-0.003^{*}	-0.004^{***}	-0.004^{***}
	(0.002)	(0.001)	(0.001)
TV Dummy $\times$ Distance to Destination	-0.004^{***}	-0.002	-0.002
	(0.001)	(0.001)	(0.001)
Distance from Contour to Origin (KM)	-0.0003	0.001	0.001
	(0.001)	(0.001)	(0.001)
Distance from Contour to Destination (KM)	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.0002)	(0.0003)	(0.0003)
Observations	4,062	4,062	4,062
Log(Population)	Yes	Yes	Yes
County % Hispanic	No	Yes	Yes
Log(Income)	No	No	Yes

Notes: County-county data with origin F.E.

Evidence of first stage

- Data from American Time Use Survey over last 15 years:
  - 210,000 person-year observations
  - Median person watches 120 minutes of TV per day
- Not the most satisfactory data:
  - County-level data is noisy for RD approach
  - Only overall TV viewership data, not SLTV

TV Viewership within the SLTV Boundary

Panel A: Minutes of TV watched	(1)	(2)	(3)	(4)
TV Dummy × Hispanic	13.610***	14.638***	8.777**	7.911**
	(3.825)	(3.831)	(3.834)	(3.829)
TV Dummy	−1.170	−1.392	4.096**	5.030**
	(1.944)	(1.945)	(1.960)	(1.958)
Observations	91,315	91,315	91,315	91,315
County % Hispanic	No	Yes	Yes	Yes
Log(Income)	No	No	Yes	Yes
Person is Migrant	No	No	No	Yes

Notes:

Schools

Schools

- Data from Department of Education's Civil Rights Data Collection in 2015:
  - 48,000 public schools in sample (unit of observation)
  - (Almost) all variables split by ethnicity
- Educational Attainment
  - Gifted Program Enrolment
  - AP Program Enrolment and Exam Passes
  - Limited English Proficiency (LEP)
- Discipline
  - Out of School Suspensions
  - Chronic Absentees
  - Ethnicity-Based Bullying/Harassment

Schools - Summary Statistics - Outcomes

	All	*No TV*	*TV*
	(1)	(2)	(3)
Panel A: Schools			
IHS(Hispanic Gifted Students)	1.988	1.262	2.380
	(1.552)	(1.238)	(1.563)
IHS(Hispanic AP Enrolment)	3.192	2.091	3.778
	(1.937)	(0.646)	(0.918)
IHS(Hispanic AP Passes)	4.087	3.497	4.181
	(0.917)	(0.646)	(0.918)
IHS(Hispanic Suspensions)	0.957	0.676	1.102
	(1.273)	(1.044)	(1.353)
IHS(Hispanic Absentees)	2.655	1.881	3.054
	(1.765)	(1.536)	(1.742)
IHS(Hispanic Limited English Proficiency)	2.915	2.113	3.331
	(2.040)	(1.820)	(2.024)
IHS(Hispanic Harassment)	0.045	0.027	0.055
	(0.273)	(0.211)	(0.299)
Observations	41,502	11,252	30,250

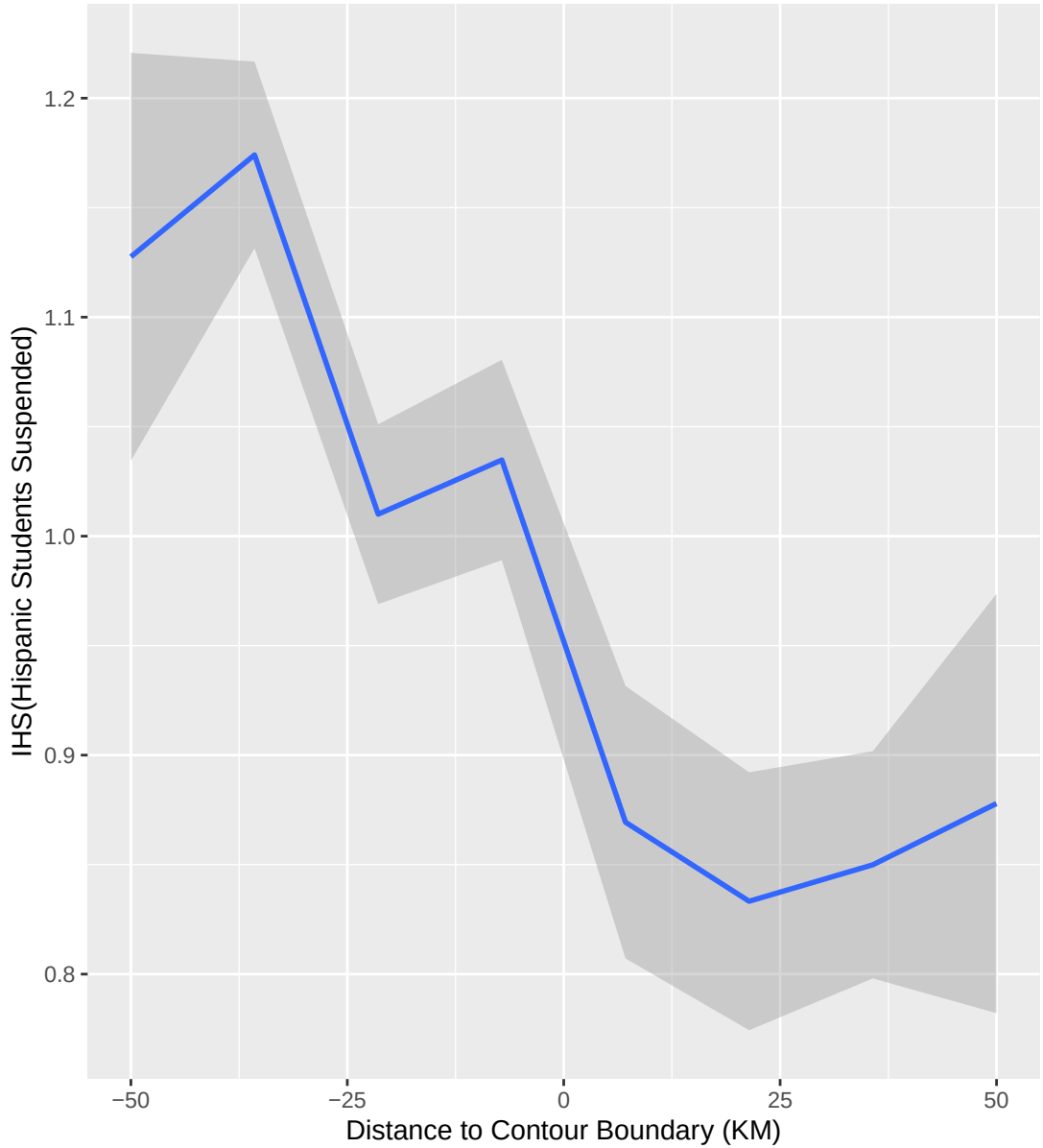
Notes: The table presents means (and standard deviations).

Schools - Summary Statistics - Controls

	All	*No TV*	*TV*
	(1)	(2)	(3)
Panel A: Schools			
Log Income	9.547	9.430	9.608
	(0.303)	(0.200)	(0.328)
Log Population	12.484	11.559	12.964
	(1.576)	(1.471)	(1.405)
Fraction County Hispanic	0.107	0.037	0.143
	(0.160)	(0.079)	(0.179)
# School Teachers	39.591	32.684	43.169
	(30.764)	(24.090)	(33.146)
# Hispanic Students	164.343	68.500	214.011
	(259.096)	(117.433)	(295.883)
# Total Students	581.524	478.166	635.086
	(482.595)	(383.924)	(518.467)
Observations	41,502	11,252	30,250

Notes: The table presents means (and standard deviations). No control is significantly different across the coverage contour at the $\alpha = .1$ level.

Main Findings - Schools



Predicted coefficients and 95% CI from regression

Effect of SLTV on Hispanic Educational Attainment

Panel A: IHS(# Hispanic Gifted Students)	(1)	(2)	(3)
TV Dummy	0.016***	0.015**	0.013**
	(0.006)	(0.006)	(0.006)
Observations	26,065	26,065	26,065
Panel B: IHS(# Hispanic Students Taking AP)			
TV Dummy	0.072***	0.051***	0.047***
	(0.016)	(0.015)	(0.015)
Observations	6,089	6,089	6,089
Panel C: IHS(# Hispanic Students Passing AP)			
TV Dummy	0.034**	0.042***	0.039***
	(0.014)	(0.013)	(0.013)
Observations	2,205	2,205	2,205
County Controls	Yes	Yes	Yes
School Size Controls	No	Yes	Yes
School Type Controls	No	No	Yes

Notes:

Effect of SLTV on Hispanic School Discipline

Panel A: IHS(Out of School Suspensions)	(1)	(2)	(3)
TV Dummy	-0.011^{**}	-0.018^{***}	-0.016^{***}
	(0.005)	(0.005)	(0.005)
Observations	40,864	40,864	40,864
Panel B: IHS(Students Chronically Absent)			
TV Dummy	-0.067^{***}	-0.073^{***}	-0.074^{***}
	(0.006)	(0.006)	(0.006)
Observations	40,869	40,869	40,869
County Controls	Yes	Yes	Yes
School Size Controls	No	Yes	Yes
School Type Controls	No	No	Yes

Notes:

Effect of SLTV on Hispanic School Identity

Panel A: IHS(Limited English Proficiency)	(1)	(2)	(3)
TV Dummy	0.040***	0.039***	0.031***
	(0.007)	(0.007)	(0.007)
Observations	40,864	40,864	40,864
Panel B: IHS(Victims of Harassment on Basis of Ethnicity)			
TV Dummy	0.003**	0.002*	0.002*
	(0.001)	(0.001)	(0.001)
Observations	40,811	40,811	40,811
County Controls	Yes	Yes	Yes
School Size Controls	No	Yes	Yes
School Type Controls	No	No	Yes

Notes:

Effect sizes

Outcome	White-Hispanic Gap	TV Regression Coeff.
Gifted students	80.0%	1.3 - 1.6 %
AP Students	39.3%	4.7 - 7.2 %
Pass AP	22.7%	3.4 - 4.2 %
Suspension	-30.0%	1.1 - 1.8 %
Absent	-12.8%	6.7 - 7.4 %

Schools: Closing Thoughts

- Robustness ▸ Robustness ▸ Spatial
- Broadly positive main effects of TV support (Gentzkow & Shapiro 2008) against mainstream findings (Gentile 2004; Zavodny 2006)
- Contrasting effect of English proficiency and harassment — consistent story of identity being strengthened

Firms

Firms

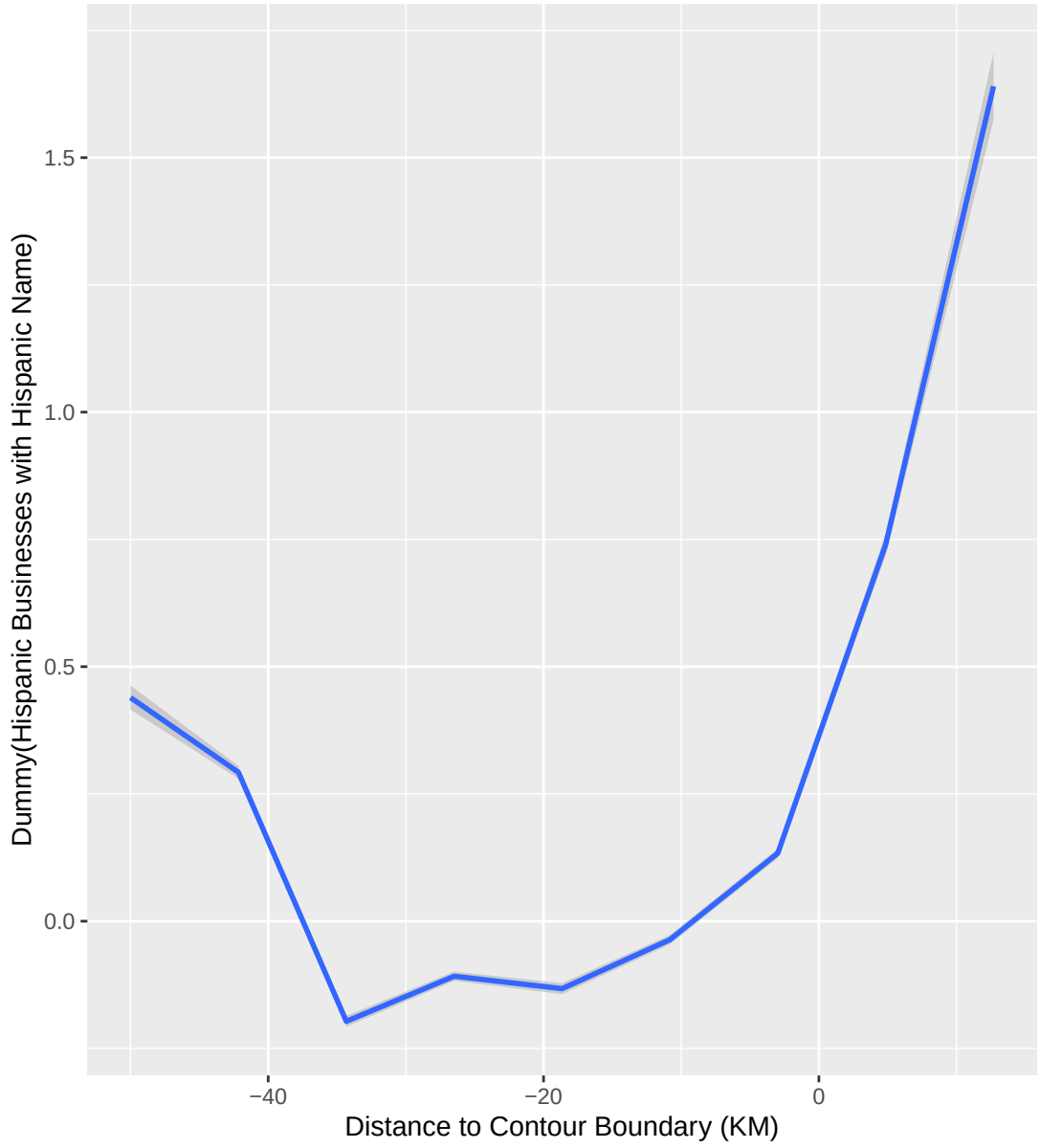
- Data from Florida's Division of Corporations
  - Why Florida? 23% Hispanic (8% US total), 11 SLTV stations (11% total) and open data
  - 146,032 firms successfully geocoded
  - Aggregate data into 2×2 KM^2 squares
- Firm Owner Name Classification
  - 'ethnicolr' — a LSTM model trained with TensorFlow on Florida voter registration data
  - Validation $>$ 85% accurate, 23.5% firm owners are Hispanic
- Firm Name Classification
  - Keyword matching on (1) references to Latin American countries, (2) top 50 most common Spanish words not in English, and optionally (3) references to common Hispanic foods
  - 1% (1.1% with food) of firms match this criteria

Firms - Summary Statistics

	All	*No TV*	*TV*
	(1)	(2)	(3)
Panel A: Firms			
IHS(Hispanic Owned Firms)	0.992	0.671	1.225
	(1.694)	(1.308)	(1.892)
Hispanic Named Firms	0.027	0.006	0.042
	(0.161)	(0.080)	(0.200)
Log Income	9.498	9.463	9.523
	(0.241)	(0.284)	(0.201)
Log Population	11.954	11.206	12.497
	(1.398)	(1.253)	(1.239)
Fraction County Hispanic	0.086	0.063	0.103
	(0.105)	(0.061)	(0.125)
Observations	23,823	10,023	13,830

Notes: The table presents means (and standard deviations). No control is significantly different across the coverage contour at the $\alpha = .1$ level.

Main Findings - Firms



Predicted coefficients and 95% CI from logit regression

Effect of SLTV on Hispanic Firm Ownership

	IHS(# Hispanic Owned Businesses)			
	(1)	(2)	(3)	(4)
TV Dummy	0.261***	0.122***	0.112***	0.132***
	(0.014)	(0.014)	(0.014)	(0.015)
TV Dummy × Distance to Boundary	0.010***	0.007***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Distance to Boundary (meters)	0.006***	0.009***	0.010***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)
Log(Population)		0.412***	0.388***	0.342***
		(0.011)	(0.012)	(0.014)
County % Hispanic			1.261***	1.414***
			(0.133)	(0.136)
Log(Income)				0.391***
				(0.070)
Observations	23,853	23,853	23,853	23,853

Notes:

Effect of SLTV on Hispanic Firm Names

	Hispanic Named Business Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
TV Dummy	0.839***	0.638***	0.637***	0.769***	0.849***	0.775***
	(0.052)	(0.066)	(0.066)	(0.071)	(0.077)	(0.071)
TV Dummy × Distance to Boundary	0.008***	0.002	0.002	0.0002	−0.0002	0.0002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Distance to Boundary (meters)	0.010**	0.021***	0.021***	0.031***	0.035***	0.031***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Log(Population)		0.957***	0.979***	0.702***	0.761***	0.701***
		(0.052)	(0.070)	(0.074)	(0.081)	(0.074)
County % Hispanic			−0.151	1.428***	1.514***	1.434***
			(0.312)	(0.367)	(0.388)	(0.368)
Log(Income)				2.350***	2.534***	2.356***
				(0.319)	(0.344)	(0.320)
Observations	23,853	23,853	23,853	23,853	23,853	23,853
Only Hispanic Owners	No	No	No	No	Yes	No
Only Non-Hispanic Owners	No	No	No	No	No	Yes

Notes: Logit regression

Firms: Closing Thoughts

- Robustness ▸ Spatial autocorrelation ▸ Vary boundary cut-off and grid size ▸ Vary controls
- Substantial effect size compared to white-Hispanic firm ownership gap (68%, log-odds ratio of **0.52**)
- Existing literature finds entrepreneurship difficult to foster (Karlan & Valdiva 2011; Gin & Mansuri 2014). We find signs of this, but no data on firm size.
- Business naming demonstrates an appeal to identity (disproportionate increase)
- Definite demand effect, supply side less clear

Conclusion

- SLTV creates a variety of effects: more firms and better academic performance (omitted: political campaign contributions)
- Underlying theme of strengthened identity that reduces inequality
- Lots of space for future work

Thank You!

Effect of SLTV on Hispanic Firm Names - Different Boundary/Grid Sizes

	Hispanic Owned & Named Business Dummy				
	(1)	(2)	(3)	(4)	(5)
TV Dummy	0.849***	0.927***	0.596***	0.624***	1.144***
	(0.077)	(0.098)	(0.118)	(0.078)	(0.076)
TV Dummy × Distance to Boundary	−0.0002	−0.002	0.042***	0.001	−0.001
	(0.002)	(0.004)	(0.010)	(0.002)	(0.002)
Distance to Boundary (meters)	0.035***	0.049***	−0.097***	0.026***	0.042***
	(0.005)	(0.012)	(0.035)	(0.005)	(0.006)
Observations	23,853	20,404	14,386	10,598	95,373
County Controls	Yes	Yes	Yes	Yes	Yes
Distance Cutoff (KM)	100	50	25	100	100
Grid Size (KM^2)	4	4	4	9	1

Notes: Logit regressions

Back

Effect of SLTV on Hispanic Educational Attainment - Robustness

	IHS(# Hispanic Students Passing AP)					
	(1)	(2)	(3)	(4)	(5)	(6)
TV Dummy	0.039***	0.049***	0.044***	0.044***	0.036***	0.032*
	(0.013)	(0.017)	(0.016)	(0.017)	(0.013)	(0.018)
TV Dummy × Distance to Boundary	0.0003	0.0001	0.001	0.001*	0.0001	0.001
	(0.0002)	(0.001)	(0.001)	(0.0004)	(0.0004)	(0.001)
Distance to Boundary (meters)	0.001	0.012***	0.006***	0.006***	0.003**	0.001
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)
Total APs Passed					0.003***	
					(0.0001)	
Observations	2,205	2,205	1,525	1,525	1,525	1,095
County/School Controls	Yes	Yes	Yes	Yes	Yes	Yes
Distance Cutoff (KM)	100	100	50	50	50	$33\frac{1}{3}$
Distance^2 Interaction	No	Yes	No	No	No	No
County F.E.	No	No	No	Yes	No	No

Notes:

Back

Effect of SLTV on Hispanic Identity - Spatial Robustness

	IHS(# Hispanic Victims of Harassment)		
	(1)	(2)	(3)
TV Dummy	0.003**	0.002***	0.003*
	(0.001)	(0.001)	(0.002)
TV Dummy × Distance to Boundary	−0.0001**	−0.0001***	−0.0001**
	(0.00002)	(0.00001)	(0.00003)
Observations	40,811	40,811	40,811
Log Likelihood		−4,304.916	−4,299.820
σ^2		0.072	0.072
Akaike Inf. Crit.		8,629.833	8,619.640
Wald Test (df = 1)		686.149***	686.981***
LR Test (df = 1)		657.312***	667.505***
County/School Controls	Yes	Yes	Yes
Model	OLS	SAR Lag	SAR Error

Notes:

Back

Effect of SLTV on Hispanic Firm Ownership - Spatial Robustness

	IHS(# Hispanic Owned Firms)		
	(1)	(2)	(3)
TV Dummy	0.122***	0.022***	0.126***
	(0.014)	(0.006)	(0.036)
Observations	23,853	23,853	23,853
Log Likelihood		−38,404	−38,440
σ^2		1.168	1.170
Akaike Inf. Crit.		76,821	76,894
Wald Test (df = 1)		65,139***	63,913***
LR Test (df = 1)		24,759***	24,687***
County Controls	Yes	Yes	Yes
Model	OLS	SAR Lag	SAR Error

Notes:

Back

Effect of SLTV on Hispanic Firm Names - Robustness

	Hispanic Owned & Named Business Dummy			
	(1)	(2)	(3)	(4)
TV Dummy	0.849***	1.071***	0.305***	.8677***
	(0.077)	(0.115)	(0.078)	(0.079)
TV Dummy × Distance to Boundary	−0.0002	−0.008	−0.003	−0.001
	(0.002)	(0.007)	(0.002)	(0.002)
Distance to Boundary (meters)	0.035***	0.123***	0.013***	0.036***
	(0.005)	(0.021)	(0.005)	(0.005)
Total Businesses			0.023***	
			(0.001)	
Observations	23,853	23,853	23,853	23,853
County Controls	Yes	Yes	Yes	Yes
Distance^2	No	Yes	No	No
No Food Names	No	No	No	Yes

Notes: Logit regressions

Back