

# Seeing is Believing: The Effect of Television on the Identity and Lives of Hispanic People

**Andrew Kao**

*Harvard*

March 2021

# Motivation

- ▶ Large literature on how TV affects behavior (Yanigazawa-Drott 2014; DellaVigna & al. 2007; Ferrara & al., 2012)
- ▶ 50% of Hispanics watch satellite or broadcast Spanish Language TV (SLTV)
- ▶ Complicated time for largest ethnic minority in the US
- ▶ Prior efforts to study Hispanic interaction with media focused on politics (Waldfoegel & al. 2009; Trujillo & al. 2012)

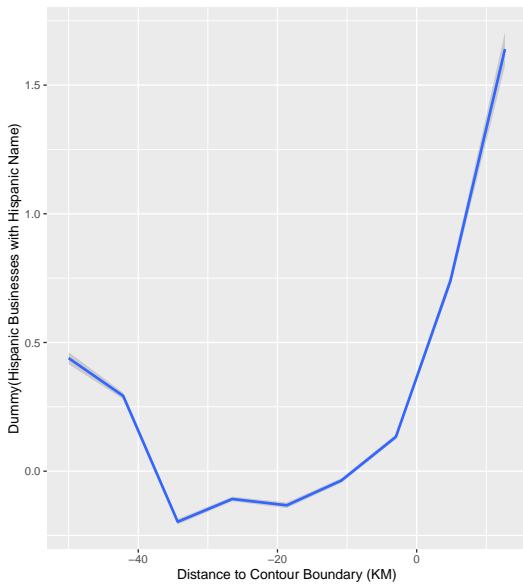
This paper examines the effect of SLTV in two ways:

**(1) How does Hispanic behavior change in firms and schools?**

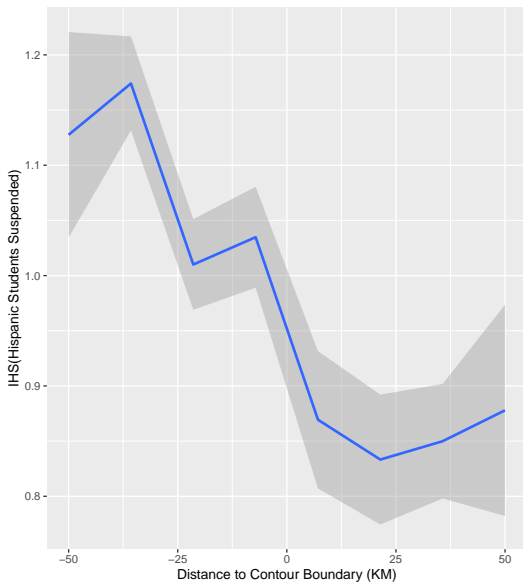
**(2) How is identity affected?**

- ▶ Identification: follow Velez & Newman (2019) and construct spatial RD arising from FCC TV signal regulation

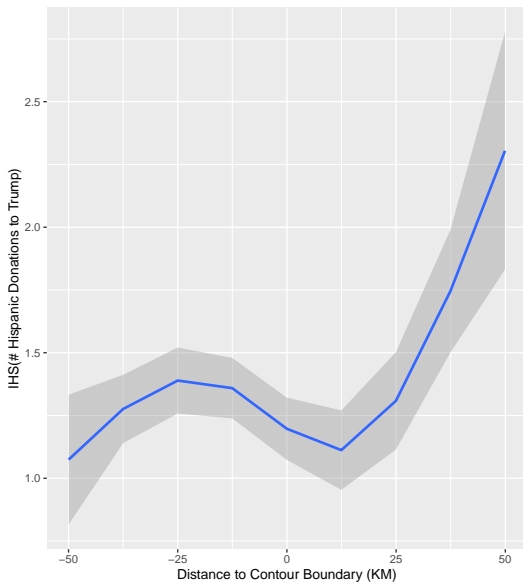
# Main Findings - Firms



# Main Findings - Schools



# Main Findings - Campaign Contributions



# 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 (Benjamin & al. 2007; Bursztyrn & al. 2015). New research on how identity is constructed and strengthened (Atkin & al. 2019; Bazzi & al. 2019)
- Supply a revealed preference demonstration of how identity can be bolstered by the media

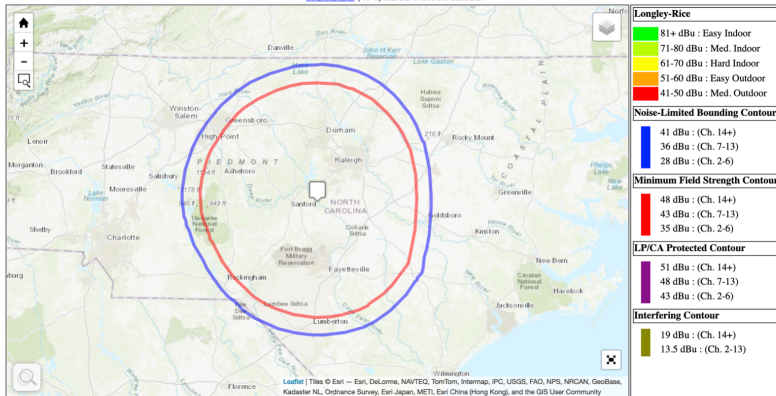
# Empirical Strategy

- ▶ 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
  - ▶ Focus on the RD, dummy for whether observation falls inside contour

# Coverage Map for TV Station WUVC-DT

## Coverage Maps

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





# Specifications

## Main Model:

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

## Spatial Autogressive:

$$Y = \beta_0 + \rho WY + \beta \mathbb{I}[InsideContour] \times Distance + \gamma X + \epsilon$$

## Spatial Error:

$$Y = \beta_0 + \beta \mathbb{I}[InsideContour] \times Distance + \gamma X + \epsilon$$

$$\epsilon = \lambda W\epsilon + \nu$$

where  $W$  is a 4 nearest neighbor/rook spatial weights matrix

# 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 and small number of campaign contributions)
  - ▶ US Census Geocoder: 80% successfully geocoded (firms and campaign contributions)
- ▶ Demographic and migration information at county level from ACS

# Selection? Migration?

	IHS(# Hispanic Migrants)		
	(1)	(2)	(3)
Panel A: Origin County Inside Contour			
Dummy: Destination Outside TV Contour	-0.387*** (0.048)	-0.286*** (0.044)	-0.280*** (0.044)
TV Dummy $\times$ Distance to Origin	-0.003** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)
TV Dummy $\times$ Distance to Destination	0.001 (0.001)	-0.002* (0.001)	-0.002 (0.001)
Distance from Contour to Origin (KM)	0.001 (0.002)	0.003* (0.002)	0.003 (0.002)
Distance from Contour to Destination (KM)	-0.001 (0.001)	0.002 (0.001)	0.002 (0.001)
Observations	8,479	8,479	8,479
Panel B: Origin County Outside Contour			
Dummy: Destination Inside TV Contour	-0.078 (0.108)	-0.123 (0.096)	-0.120 (0.096)
TV Dummy $\times$ Distance to Origin	-0.003* (0.002)	-0.004*** (0.001)	-0.004*** (0.001)
TV Dummy $\times$ Distance to Destination	-0.004*** (0.001)	-0.002 (0.001)	-0.002 (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*** (0.0002)	-0.001*** (0.0003)	-0.001*** (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.

# 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  $2 \times 2$  KM<sup>2</sup> 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

	<i>All</i>	<i>No TV</i>	<i>TV</i>
	(1)	(2)	(3)
Panel A: Firms			
IHS(Hispanic Owned Firms)	0.992 (1.694)	0.671 (1.308)	1.225 (1.892)
Hispanic Named Firms	0.027 (0.161)	0.006 (0.080)	0.042 (0.200)
Log Income	9.498 (0.241)	9.463 (0.284)	9.523 (0.201)
Log Population	11.954 (1.398)	11.206 (1.253)	12.497 (1.239)
Fraction County Hispanic	0.086 (0.105)	0.063 (0.061)	0.103 (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.

# Effect of SLTV on Hispanic Firm Ownership

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

*Notes:*

# Effect of SLTV on Hispanic Firm Ownership - Spatial Robustness

	<i>IHS(# Hispanic Owned Firms)</i>		
	(1)	(2)	(3)
TV Dummy	0.122*** (0.014)	0.022*** (0.006)	0.126*** (0.036)
Observations	23,853	23,853	23,853
Log Likelihood		-38,404	-38,440
$\sigma^2$		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:*

# Effect of SLTV on Hispanic Firm Names

	<i>IHS(# Hispanic Owned Firms)</i>		
	(1)	(2)	(3)
TV Dummy	0.122*** (0.014)	0.022*** (0.006)	0.126*** (0.036)
Observations	23,853	23,853	23,853
Log Likelihood		-38,404	-38,440
$\sigma^2$		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:*



# Effect of SLTV on Hispanic Firm Names

	<i>Hispanic Named Business Dummy</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
TV Dummy	0.839*** (0.052)	0.638*** (0.066)	0.637*** (0.066)	0.769*** (0.071)	0.849*** (0.077)	0.775*** (0.071)
TV Dummy $\times$ Distance to Boundary	0.008*** (0.002)	0.002 (0.002)	0.002 (0.002)	0.0002 (0.002)	-0.0002 (0.002)	0.0002 (0.002)
Distance to Boundary (meters)	0.010** (0.004)	0.021*** (0.004)	0.021*** (0.005)	0.031*** (0.005)	0.035*** (0.005)	0.031*** (0.005)
Log(Population)		0.957*** (0.052)	0.979*** (0.070)	0.702*** (0.074)	0.761*** (0.081)	0.701*** (0.074)
County % Hispanic			-0.151 (0.312)	1.428*** (0.367)	1.514*** (0.388)	1.434*** (0.368)
Log(Income)				2.350*** (0.319)	2.534*** (0.344)	2.356*** (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

# Effect of SLTV on Hispanic Firm Names - Robustness

	<i>Hispanic Owned &amp; Named Business Dummy</i>			
	(1)	(2)	(3)	(4)
TV Dummy	0.849*** (0.077)	1.071*** (0.115)	0.305*** (0.078)	.8677*** (0.079)
TV Dummy $\times$ Distance to Boundary	-0.0002 (0.002)	-0.008 (0.007)	-0.003 (0.002)	-0.001 (0.002)
Distance to Boundary (meters)	0.035*** (0.005)	0.123*** (0.021)	0.013*** (0.005)	0.036*** (0.005)
Total Businesses			0.023*** (0.001)	
Observations	23,853	23,853	23,853	23,853
County Controls	Yes	Yes	Yes	Yes
Distance <sup>2</sup>	No	Yes	No	No
No Food Names	No	No	No	Yes

*Notes:* Logit regressions

► Robustness: Vary Boundary Cut-Off and Grid Size

# Firms: Closing Thoughts

- ▶ Existing literature finds entrepreneurship difficult to foster (Karlan & Valdiva 2011; Gine & 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

# 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

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

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

# Schools - Summary Statistics - Controls

	<i>All</i>	<i>No TV</i>	<i>TV</i>
	(1)	(2)	(3)
Panel A: Schools			
Log Income	9.547 (0.303)	9.430 (0.200)	9.608 (0.328)
Log Population	12.484 (1.576)	11.559 (1.471)	12.964 (1.405)
Fraction County Hispanic	0.107 (0.160)	0.037 (0.079)	0.143 (0.179)
# School Teachers	39.591 (30.764)	32.684 (24.090)	43.169 (33.146)
# Hispanic Students	164.343 (259.096)	68.500 (117.433)	214.011 (295.883)
# Total Students	581.524 (482.595)	478.166 (383.924)	635.086 (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.

# Effect of SLTV on Hispanic Educational Attainment

Panel A: IHS(# Hispanic Gifted Students)			
	(1)	(2)	(3)
TV Dummy	0.016*** (0.006)	0.015** (0.006)	0.013** (0.006)
Observations	26,065	26,065	26,065
Panel B: IHS(# Hispanic Students Taking AP)			
TV Dummy	0.072*** (0.016)	0.051*** (0.015)	0.047*** (0.015)
Observations	6,089	6,089	6,089
Panel C: IHS(# Hispanic Students Passing AP)			
TV Dummy	0.034** (0.014)	0.042*** (0.013)	0.039*** (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** (0.005)	-0.018*** (0.005)	-0.016*** (0.005)
Observations	40,864	40,864	40,864
Panel B: IHS(Students Chronically Absent)			
TV Dummy	-0.067*** (0.006)	-0.073*** (0.006)	-0.074*** (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*** (0.007)	0.039*** (0.007)	0.031*** (0.007)
Observations	40,864	40,864	40,864
Panel B: IHS(Victims of Harassment)			
TV Dummy	0.003** (0.001)	0.002* (0.001)	0.002* (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:*

# Schools: Closing Thoughts

- ▶ Robustness ▶ Robustness ▶ Spatial
- ▶ Broadly prosocial main effects support (Gentzkow & Shapiro 2008) against mainstream findings (Gentile 2004; Zavodny 2006)
- ▶ Contrasting effect within identity — consistent story of identity being strengthened

# 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

	<i>Hispanic Owned &amp; Named Business Dummy</i>				
	(1)	(2)	(3)	(4)	(5)
TV Dummy	0.849*** (0.077)	0.927*** (0.098)	0.596*** (0.118)	0.624*** (0.078)	1.144*** (0.076)
TV Dummy $\times$ Distance to Boundary	-0.0002 (0.002)	-0.002 (0.004)	0.042*** (0.010)	0.001 (0.002)	-0.001 (0.002)
Distance to Boundary (meters)	0.035*** (0.005)	0.049*** (0.012)	-0.097*** (0.035)	0.026*** (0.005)	0.042*** (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 <sup>2</sup> )	4	4	4	9	1

Notes: Logit regressions

▶ Back

## Effect of SLTV on Hispanic Educational Attainment - Robustness

	<i>IHS(# Hispanic Students Passing AP)</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
TV Dummy	0.039*** (0.013)	0.049*** (0.017)	0.044*** (0.016)	0.044*** (0.017)	0.036*** (0.013)	0.032*** (0.018)
TV Dummy $\times$ Distance to Boundary	0.0003 (0.0002)	0.0001 (0.001)	0.001 (0.001)	0.001* (0.0004)	0.0001 (0.0004)	0.001 (0.001)
Distance to Boundary (meters)	0.001 (0.001)	0.012*** (0.003)	0.006*** (0.002)	0.006*** (0.002)	0.003** (0.002)	0.001 (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 <sup>2</sup> Interaction	No	Yes	No	No	No	No
County F.E.	No	No	No	Yes	No	No

Notes:

## Effect of SLTV on Hispanic Identity - Spatial Robustness

	<i>IHS(# Hispanic Victims of Harassment)</i>		
	(1)	(2)	(3)
TV Dummy	0.003** (0.001)	0.002*** (0.001)	0.003* (0.002)
TV Dummy $\times$ Distance to Boundary	-0.0001** (0.00002)	-0.0001*** (0.00001)	-0.0001** (0.00003)
Observations	40,811	40,811	40,811
Log Likelihood		-4,304.916	-4,299.820
$\sigma^2$		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:*