

Seeing is Believing: Identity, Inequality, and the Impact of Television on the Hispanic Achievement Gap

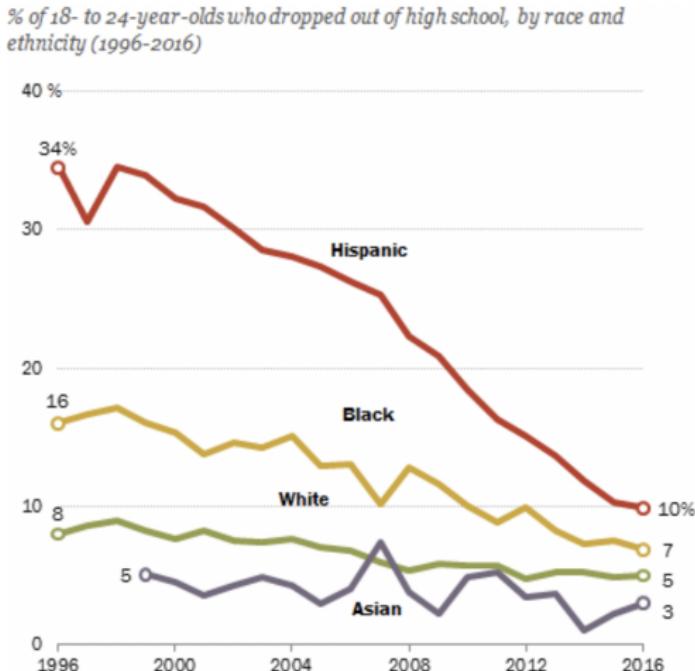
Andrew Kao
Harvard

Lab talk — Feb. 2023

Advice I'd like to get:

1. Should I keep working on this project or kill it?
Could kill it because (1) not credible or (2) too much effort required to make it publishable
2. Standard feedback (what's interesting to explore, weaknesses, etc.)

Motivation (I): the Hispanic achievement gap is large



Note: Civilian noninstitutionalized population. Blacks and Asians include the Hispanic portions of those groups. Whites include only non-Hispanics. Hispanics are of any race.
Source: U.S. Census Bureau October Current Population Survey.

PEW RESEARCH CENTER

Motivation (II): Americans *really* love TV

Outside of Sleeping, Americans Spend Most of Their Time Watching Television

Television is the most popular leisure activity in the U.S. These states tend to watch the most.

By [Kaia Hubbard](#) | July 22, 202

America's Addiction to Television Is Not Normal

Binge-watching is far from being a US-only phenomenon, but the US is way ahead of the rest of the world in its TV addiction.

By John McDuling, Quartz and Quartz

Americans Spent More Time Watching TV Than Working After Pandemic Hit in 2020

By [Terence P. Jeffrey](#) | October 15, 2021 | 4:58pm EDT

How does Spanish Language TV
affect Hispanic educational
outcomes?

This project:

Show that SLTV reduces the Hispanic achievement gap in public schools:

- ▶ Identification: difference-in-discontinuities design
- ▶ Gap vs. whites and Asians in SAT/ACTs taken, calculus courses taken, AP exams passed, etc. **shrinks** with SLTV
- ▶ However, the gap vs. whites and Asians **rises** when looking at English proficiency

How to reconcile this?

How to reconcile this?

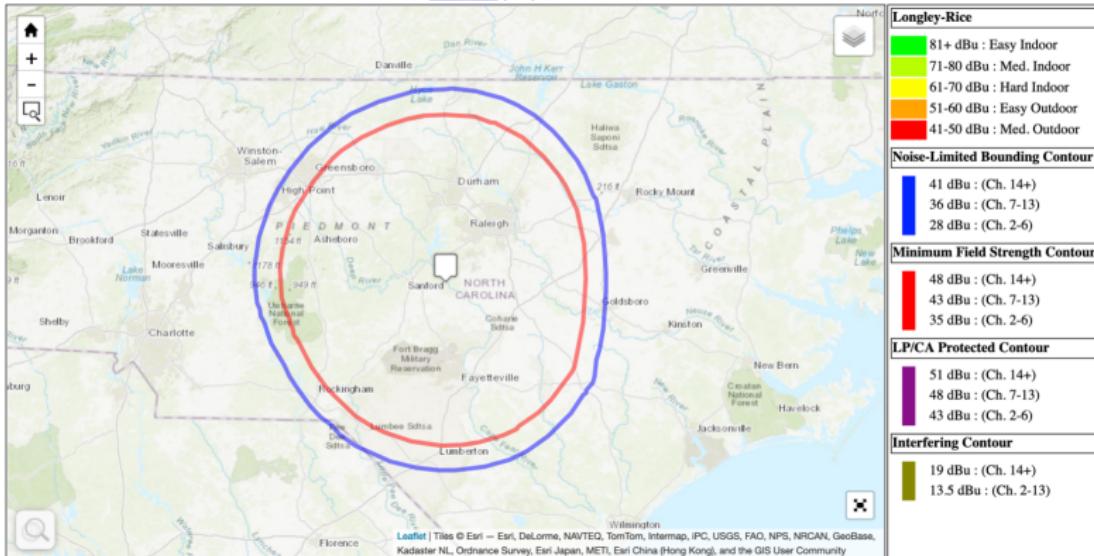
Propose an **identity** mechanism. Four strands of evidence:

1. More bullying on the basis of ethnicity (but not gender)
2. Hispanics perform better where SLTV focuses more on Hispanic identity (but not on education)
3. Hispanics with SLTV visit Hispanic branded establishments more (but not Brazilian branded ones)
4. Counties with SLTV are more socially connected to LatAm

Coverage Map for TV Station WUVC-DT

Coverage Maps

[WUVC-DT \(40-1\) BLCDT-20060912ACZ](#)



Empirical Strategy

- ▶ Construct fuzzy spatial RD arising from FCC TV signal regulation (*OET Bulletin 69*)
 - ▶ TV stations protected from interference only within certain coverage contour areas. Keep observations within 100 KM of the contour boundary
 - ▶ Follow Velez & Newman (2019), expand from 2 counties to entire US
 - ▶ Spanish Language TV: Isolate causal effect on SLTV on Hispanic communities

Empirical Strategy

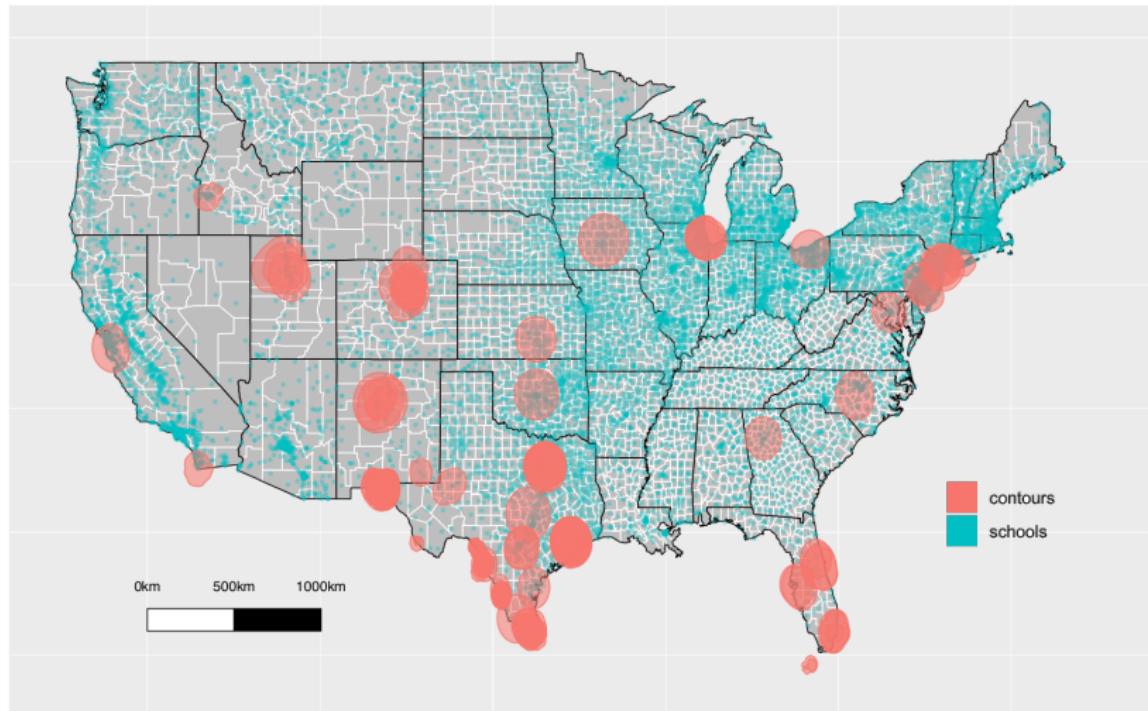
- ▶ Construct fuzzy spatial RD arising from FCC TV signal regulation (*OET Bulletin 69*)
 - ▶ TV stations protected from interference only within certain coverage contour areas. Keep observations within 100 KM of the contour boundary
 - ▶ Follow Velez & Newman (2019), expand from 2 counties to entire US
 - ▶ Spanish Language TV: Isolate causal effect on SLTV on Hispanic communities
- ▶ Compare against outcomes among Asians
 - ▶ Less likely to identify as Hispanic (or watch SLTV)
 - ▶ Combine RD with Asian 'control' for difference in discontinuities

Empirical Specification

$$y_{i,j} = \beta \mathbb{I}[InsideContour_{i,j}] \times \mathbb{I}[Hispanic_{i,j}] + \gamma_k + \delta X_i + \epsilon_{i,j}$$

where $y_{i,j}$ is an outcome for observation i (which may be an individual, school, or establishment) under demographic category $j \in \{\text{Hispanic, not Hispanic}\}$, γ_k is fixed effect for school district k , and X is a vector of controls for the observation.

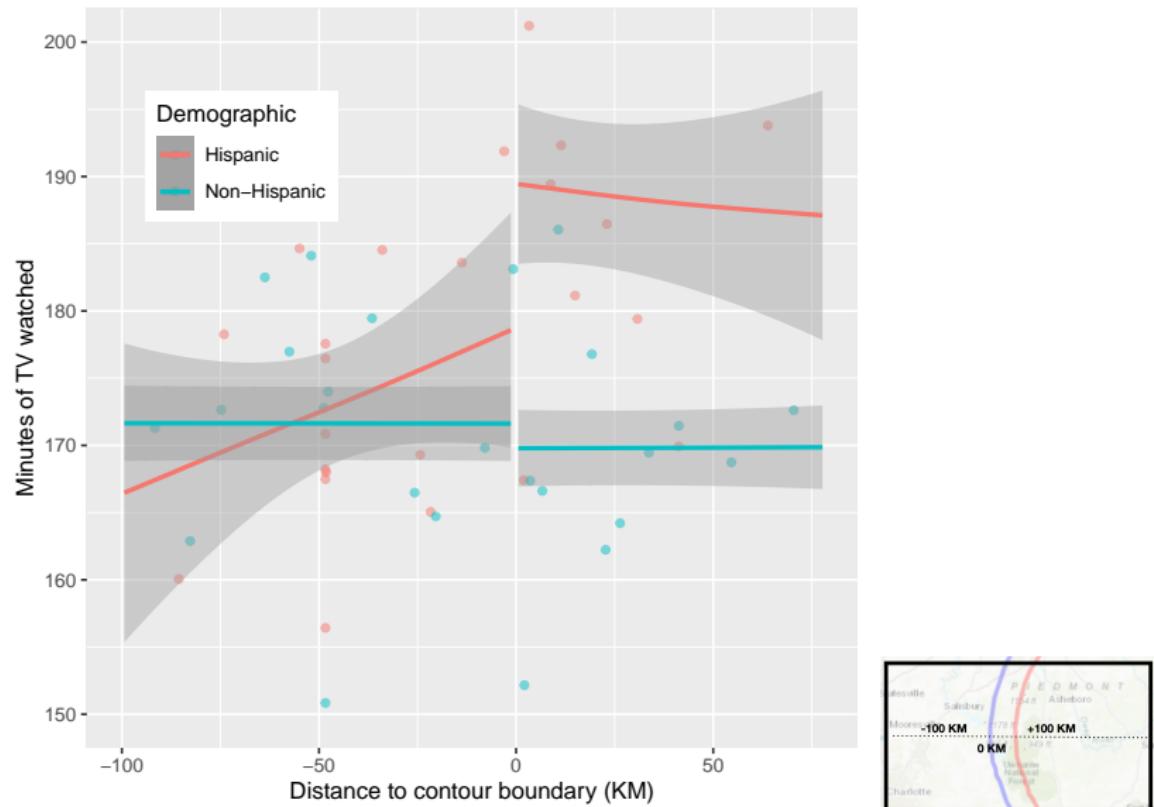
SLTV coverage and public schools



Data

- ▶ Instrument:
 - ▶ Identify 100 Spanish Language TV stations across the US from **TMS**
 - ▶ Station contours and other station data from the **FCC**
- ▶ **American Time Use Survey** over last 15 years:
 - ▶ 210,000 person-year observations
 - ▶ Average person watches 170 minutes of TV per day
- ▶ Department of Education's **Civil Rights Data Collection** in 2015:
 - ▶ 48,000 public schools in sample (unit of observation)
 - ▶ Data on academic outcomes (SAT/ACTs taken, AP exams passed, etc.) & other school data
- ▶ Other measures of identity:
 - ▶ TV transcript data from **archive.org**
 - ▶ Foot-traffic data from **Safegraph**
 - ▶ Social connectedness data from **Facebook**

TV Viewership across the SLTV Boundary



Negative distances are outside coverage contour (no SLTV).

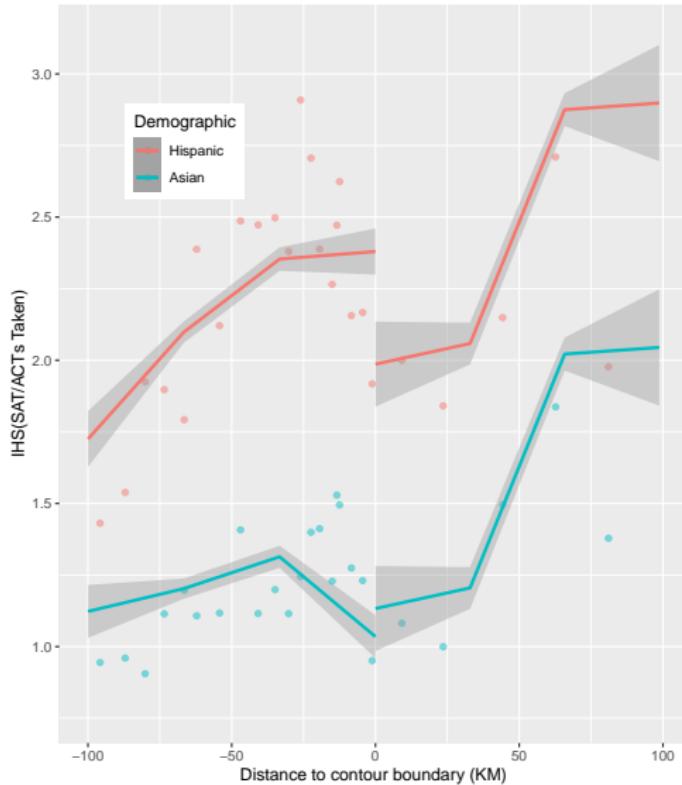
Effect of SLTV on the Hispanic achievement gap

Effect of SLTV on Hispanic vs. Asian academic achievement

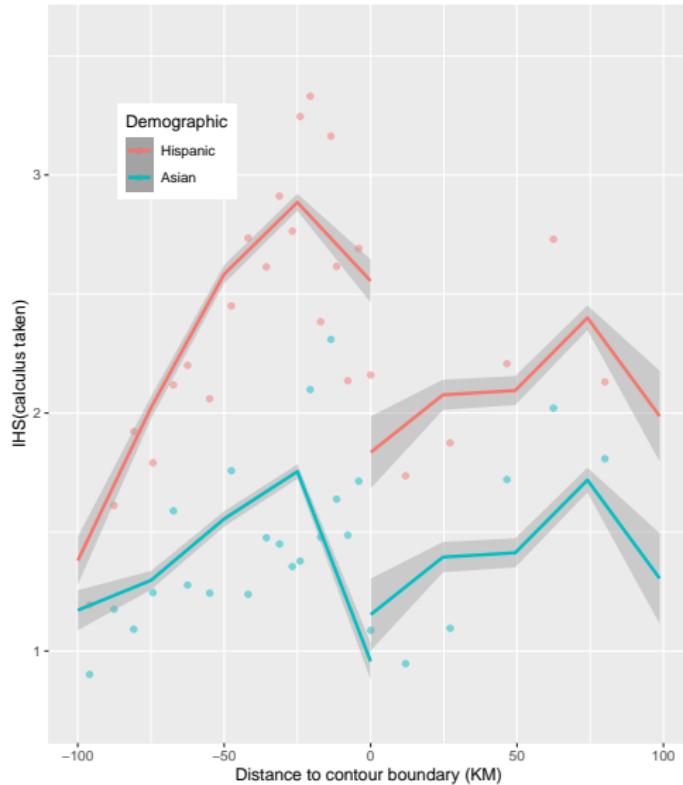
	(1)	(2)	(3)
Panel A: IHS(SAT/ACTs taken)			
TV dummy × Hispanic	0.1598*** (0.0264)	0.1598*** (0.0264)	0.1598*** (0.0264)
Panel B: IHS(calculus taken)			
TV dummy × Hispanic	0.2718*** (0.0369)	0.2718*** (0.0369)	0.2718*** (0.0369)
Panel C: IHS(APs passed)			
TV dummy × Hispanic	0.0964*** (0.0346)	0.0966*** (0.0353)	0.0972*** (0.0360)
# Hispanic, Asian students	Yes	Yes	Yes
School size controls	No	Yes	Yes
School type controls	No	No	Yes

Notes: School district fixed effects are always included. Standard errors are clustered at the school district level.

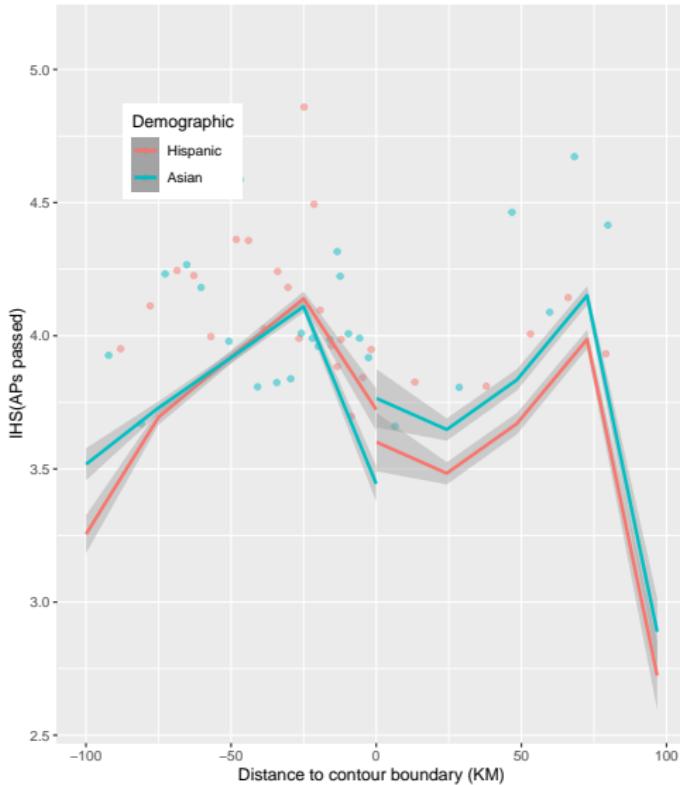
SAT/ACTs taken



Calculus taken



Pass AP exams



Effect sizes: how much is inequality reduced?

	Gap vs. Asian	Gap after SLTV
	(1)	(2)
SAT/ACTs taken	46.8%	38.3%
Calculus taken	53.6%	41.0%
APs passed	72.3%	69.6%
Gifted students	60.5%	51.0%
Advanced math taken	45.3%	31.7%
Biology taken	5.6%	-18.9%
Physics taken	43.7%	26.2%
Chemistry taken	27.7%	6.7%

So should we stick kids in front of a TV instead of sending them to school?

Exploring the identity mechanism

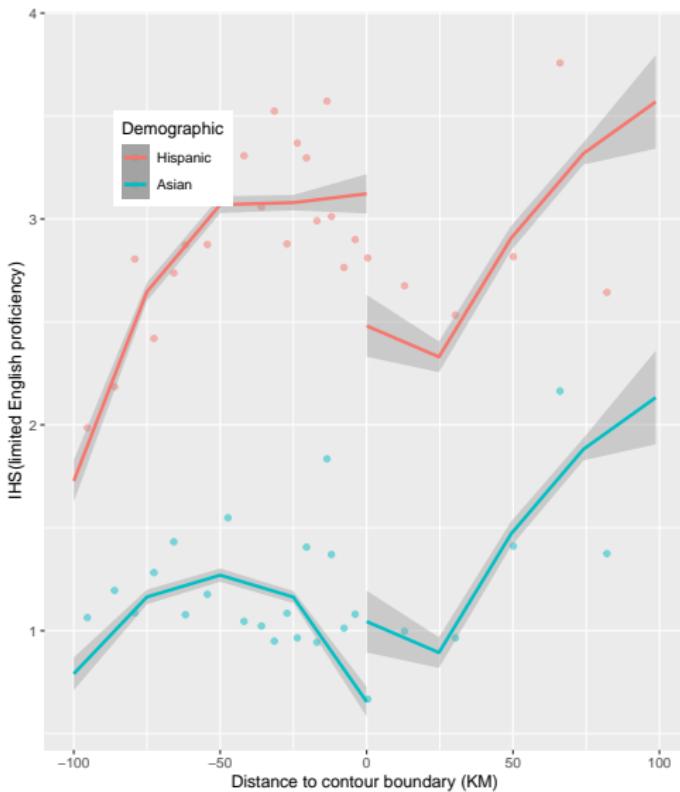
Effect of SLTV on Hispanic vs. Asian identity outcomes

	(1)	(2)	(3)
Panel A: IHS(limited English proficiency)			
TV dummy × Hispanic	0.3042*** (0.0379)	0.3042*** (0.0379)	0.3042*** (0.0379)
Panel B: IHS(bullied based on ethnicity)			
TV dummy × Hispanic	0.0015* (0.0009)	0.0015* (0.0009)	0.0015* (0.0009)
# Hispanic, Asian students	Yes	Yes	Yes
School size controls	No	Yes	Yes
School type controls	No	No	Yes

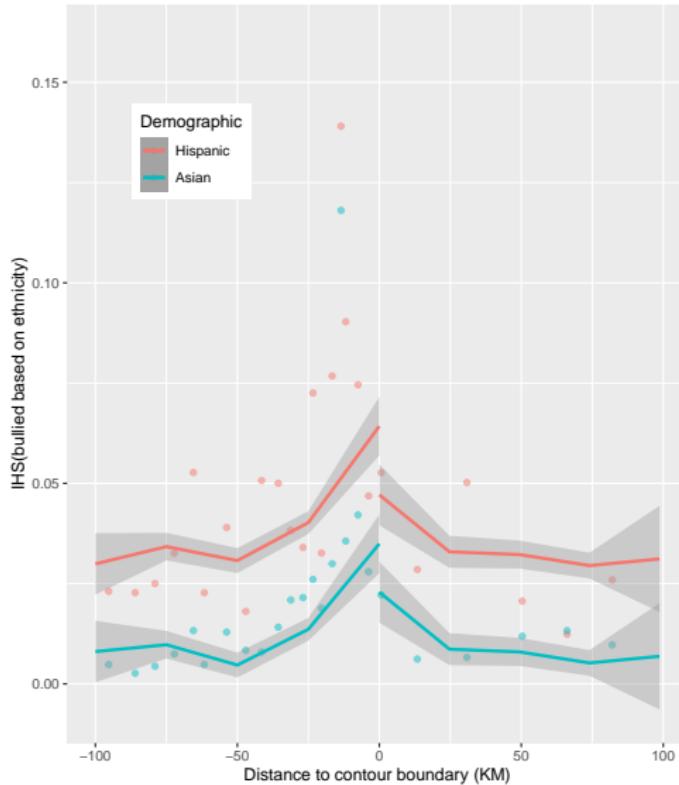
Notes: School district fixed effects are always included.
Standard errors are clustered at the school district level.

- Disability and gender-based bullying placebo

Limited English Proficiency



Bullied based on race/ethnicity



The content of SLTV programs

- ▶ Data from archive.org's TV transcript database (2005 - 2015)
 - ▶ Use keyword matching to code content of television programs
 - ▶ Variation at the television network level
- ▶ Test three different mechanisms:
 - ▶ **Identity:** 10.8% of programs relate to Latin America (vs. sports/weather/local news translated into Spanish etc.)
 - ▶ **Education:** 15% of programs that mention schools
 - ▶ **Role models:** 5.0% of programs with good role models for children/adolescents (mostly telenovelas)

Differential effect of SLTV by program content

	(1)	(2)	(3)
Panel A: IHS(SAT/ACTs taken)			
TV \times Hispanic \times % programs on identity	2.313** (0.943)		
TV \times Hispanic \times % programs on education		-0.516 (0.626)	
TV \times Hispanic \times % programs with role models			-2.085 (2.151)
# Hispanic, Asian students	Yes	Yes	Yes
School size controls	No	Yes	Yes
School type controls	No	No	Yes

Notes: School district fixed effects are always included. Standard errors are clustered at the school district level. See effect for:

▶ Calculus

▶ AP exams

More on identity

- ▶ Hispanics visit more Hispanic-branded establishments (restaurants, recreation businesses) when they have access to SLTV ▶ Data and table
 - ▶ However, they are no more likely to visit Brazilian establishments (or Japanese, or Cajun/Creole etc.)
- ▶ Counties with SLTV are more socially connected with Latin America ▶ Data and table
 - ▶ However, they are not more connected to Brazil (or the rest of the world)

Contribution

- ▶ Gentzkow & Shapiro 2008 shows how English language TV benefits students: English acquisition/cognitive channel
- Provide evidence on non-cognitive, *identity* based channel

Contribution

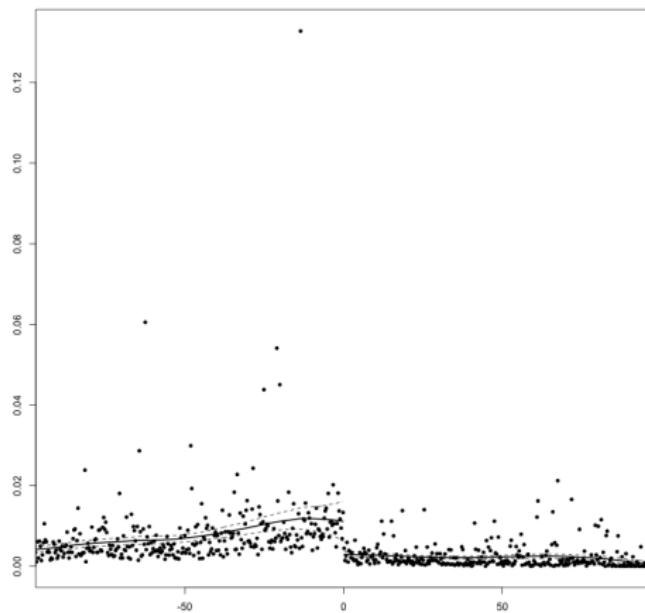
- ▶ Gentzkow & Shapiro 2008 shows how English language TV benefits students: English acquisition/cognitive channel
 - Provide evidence on non-cognitive, *identity* based channel
- ▶ Existing research that shows identity is a powerful mechanism driving meaningful outcomes (Benjamin & al. 2007; Bursztyn & al. 2015). New research on how identity is constructed and strengthened (Atkin & al. 2019; Bazzi & al. 2019)
- ▶ In much of the education lit. a salient minority identity is bad because of reasons like stereotype threat (Spencer, Logel, & Davies 2016)
 - Show how identity can be bolstered by the media and how it can help reduce inequality

Conclusion

- ▶ Hopefully persuaded you that an identity mechanism matters for Hispanic educational achievement
 - ▶ But there could also be other important ones!
 - ▶ TV appears to be one way to operationalise the identity mechanism in schools, what are others?
- ▶ Many ways that identity mechanism itself could operate (meta-mechanisms):
 - ▶ Stronger ties abroad
 - ▶ Self-confidence from representation on screen
 - ▶ Stronger in-group ties within school community
 - ▶ Greater connection with parents and support network
 - ▶ Recognise relative privilege vs. countries of origin & raise perceived value of education
 - ▶ More engagement and intellectual stimulation

Flaws

1. McCrary test fails with $p = 0.000$



...but typical selection story would expect bunching on the other side.

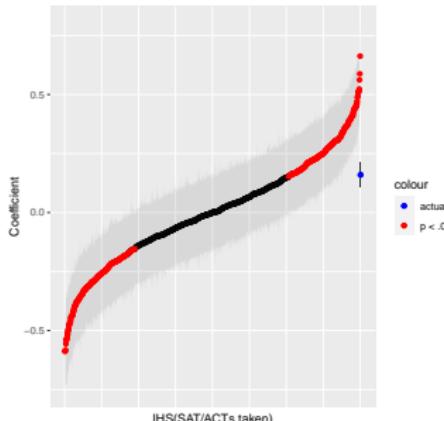
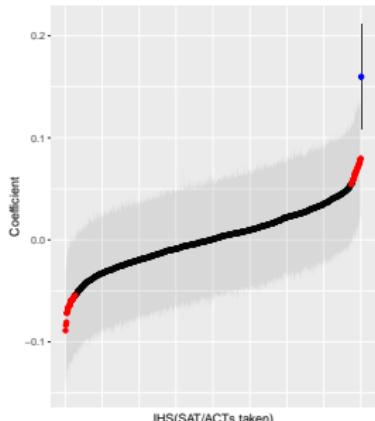
Flaws

1. McCrary test fails
2. Optimal bandwidth
 - ▶ Imbens-Kalyanamaran: 2.68 KM, not powered for this (but designed for sharp, not fuzzy)
 - ▶ All other estimators (Cattaneo, Calonico et al.) seem to have more reasonable bandwidth at about 30 KM, where my results hold

... but Lukas Althoff didn't do either and got a flyout at Harvard, so do these matter?

Flaws

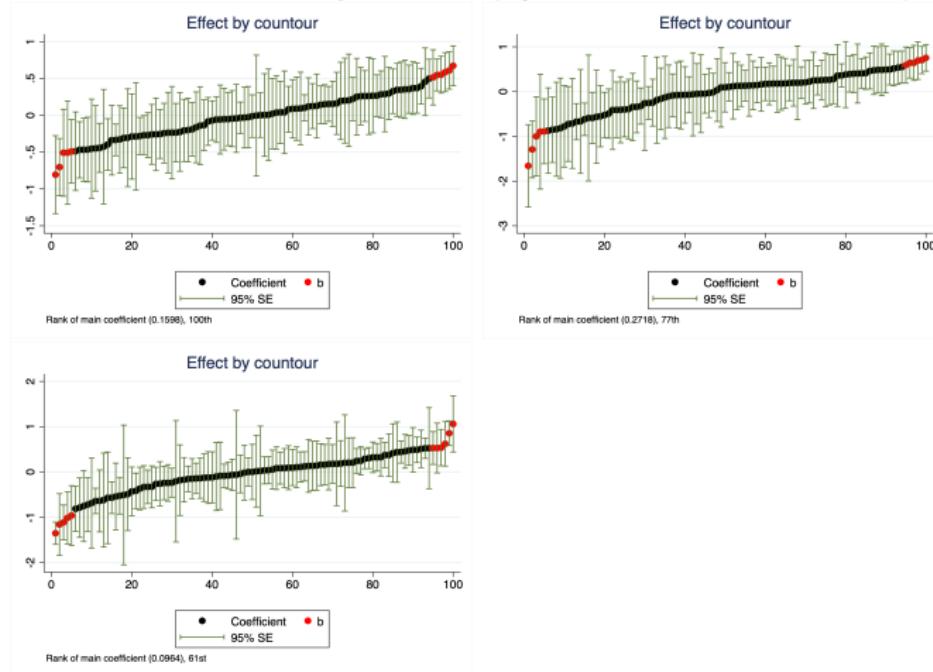
1. McCrary test fails
2. Optimal bandwidth
3. Placebo tests: Left randomly assigns TV at the school level, right randomly draws contours



Note: I have school district FEs (so should be comparing similar schools) and cluster SE at the school district level. If I allow for Conley errors (spatial autocorrelation):

Flaws

If I allow for Conley errors (spatial autocorrelation):

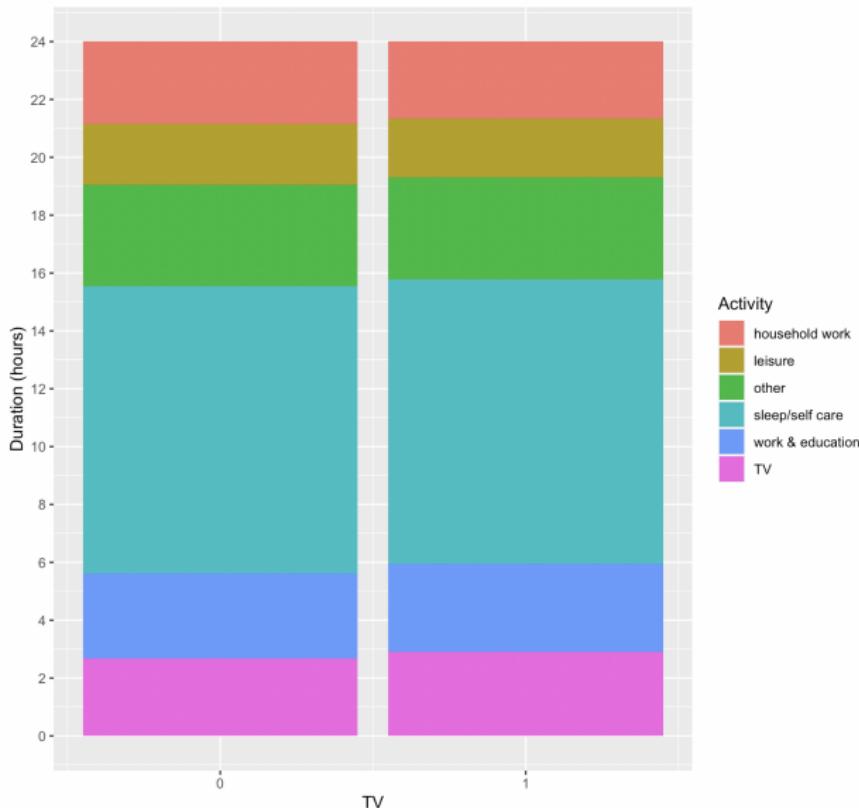


Extensions

1. ANES data for attitudes: are Hispanics more or less hardworking, etc. (test value transmission as mechanism)
 - ▶ IRB timeline has stretched over a year with no end in sight...
2. Test ties abroad: Remittances back to LatAm?
3. Self-confidence/support for schools/family ties and other mechanisms: social media + deep learning...

Thank You!

TV viewership across the SLTV boundary



Hispanics with and without SLTV ▶ Back

Differential effect of SLTV by program content

	(1)	(2)	(3)
Panel B: IHS(calculus taken)			
TV \times Hispanic \times % programs on identity	2.788*** (1.034)		
TV \times Hispanic \times % programs on education		0.829 (0.666)	
TV \times Hispanic \times % programs with role models			1.616 (2.463)
# Hispanic, Asian students	Yes	Yes	Yes
School size controls	No	Yes	Yes
School type controls	No	No	Yes

Notes: School district fixed effects are always included. Standard errors are clustered at the school district level.

[Back](#)

Differential effect of SLTV by program content

	(1)	(2)	(3)
Panel C: IHS(APs passed)			
TV \times Hispanic \times % programs on identity	1.721 (1.280)		
TV \times Hispanic \times % programs on education		0.903 (0.922)	
TV \times Hispanic \times % programs with role models			-1.184 (2.989)
# Hispanic, Asian students	Yes	Yes	Yes
School size controls	No	Yes	Yes
School type controls	No	No	Yes

Notes: School district fixed effects are always included. Standard errors are clustered at the school district level.

[Back](#)

Effect of SLTV on foot traffic

	(1)	(2)	(3)
Panel A: IHS(IDEA (disability) students)			
TV dummy \times Hispanic	0.0318 (0.0338)	0.0325 (0.0339)	0.0318 (0.0338)
Panel B: IHS(bullied based on sex)			
TV dummy \times Hispanic	0.0090 (0.0056)	0.0088 (0.0055)	0.0088 (0.0055)
# Hispanic, Asian students	Yes	Yes	Yes
School size controls	No	Yes	Yes
School type controls	No	No	Yes

Notes: School district fixed effects are always included. Standard errors are clustered at the school district level.

[Back](#)

Data: foot traffic

- ▶ Safegraph foot traffic data in 2019 to 136,000 establishments across the US
 - ▶ Restaurants are coded by Safegraph into different types of cuisine (11.6% are Hispanic)
 - ▶ Other recreational establishments are manually classified using keyword matching (10.7% are Hispanic)
 - ▶ Use census data to impute identity of visitors
- ▶ Run regressions at the establishment-visitor demographic level to see what kind of places Hispanics are more likely to visit
 - ▶ Main focus is on Hispanic-branded establishments
 - ▶ Three placebos: Brazilian, Japanese, and Creole/Cajun

Effect of SLTV on foot traffic

	<i>IHS(visitors to location)</i>			
	(1)	(2)	(3)	(4)
Panel A.1: Restaurants — Hispanic establishment indicator				
TV × Hispanic × Hispanic food	0.872*** (0.062)	0.872*** (0.062)	0.872*** (0.062)	0.872*** (0.062)
Panel B.1: Recreation — Hispanic establishment indicator				
TV × Hispanic × Hispanic brand	0.569*** (0.137)	0.569*** (0.137)	0.569*** (0.137)	0.569*** (0.137)
County log(income)	Yes	Yes	Yes	Yes
County % Hispanic	No	Yes	Yes	Yes
County log(pop.)	No	No	Yes	Yes
County FE	No	No	No	Yes
NAICS code FE	No	No	No	Yes

Notes: Standard errors are clustered at the county level. See placebos for:
► Brazilian ► Japanese ► Creole/Cajun establishments. ► Back

Data: Social Connectedness

- ▶ Facebook Social Connectedness Data from 2020
 - ▶ County-country pairs coded by $\frac{FBConnections_{ij}}{FBUsers_i \times FBUsers_j}$
 - ▶ Captures relative strength of connection between US county and foreign country
 - ▶ Connectedness with other countries ranges from 0.02 to 381, mean value 9.388

Effect of SLTV on connection to Latin America

	<i>SCI index</i>		
	(1)	(2)	(3)
Panel A: Latin America vs. rest of world			
TV dummy × Latin America	22.023*** (6.837)	22.023*** (6.838)	22.023*** (6.839)
Panel B: Latin America vs. Brazil			
TV dummy × Latin America	19.703*** (6.219)	19.703*** (6.220)	19.703*** (6.221)
County log(income)	Yes	Yes	Yes
County % Hispanic	No	Yes	Yes
County log(pop.)	No	No	Yes

Notes: Standard errors are clustered at the state level. [Back](#)

Effect of SLTV on foot traffic to Brazilian establishments

	<i>IHS(visitors to location)</i>			
	(1)	(2)	(3)	(4)
Panel A.2: Restaurants — Brazilian establishment indicator				
Hispanic × TV × Brazilian food	0.058 (0.241)	0.058 (0.241)	0.058 (0.241)	0.058 (0.241)
Panel B.2: Recreation — Brazilian establishment indicator				
Hispanic × TV × Brazilian brand	0.328 (0.598)	0.328 (0.598)	0.328 (0.599)	0.328 (0.610)
County log(income)	Yes	Yes	Yes	Yes
County % Hispanic	No	Yes	Yes	Yes
County log(pop.)	No	No	Yes	Yes
County FE	No	No	No	Yes
NAICS code FE	No	No	No	Yes

Notes: Standard errors are clustered at the county level. [Back](#)

Effect of SLTV on foot traffic to Japanese establishments

	<i>IHS(visitors to location)</i>			
	(1)	(2)	(3)	(4)
Panel A.3: Restaurants — Japanese establishment indicator				
TV × Hispanic × Japanese food	0.010 (0.067)	0.010 (0.067)	0.010 (0.067)	0.010 (0.067)
Panel B.3: Recreation — Japanese establishment indicator				
TV × Hispanic × Japanese brand	0.702 (0.528)	0.702 (0.528)	0.702 (0.528)	0.702 (0.528)
County log(income)	Yes	Yes	Yes	Yes
County % Hispanic	No	Yes	Yes	Yes
County log(pop.)	No	No	Yes	Yes
County FE	No	No	No	Yes
NAICS code FE	No	No	No	Yes

Notes: Standard errors are clustered at the county level.



Effect of SLTV on foot traffic to Cajun/Creole establishments

	<i>IHS(visitors to location)</i>			
	(1)	(2)	(3)	(4)
Panel A.4: Restaurants — Cajun and Creole establishment indicator				
TV × Hispanic × Cajun and Creole food	0.174 (0.196)	0.174 (0.196)	0.174 (0.196)	0.174 (0.196)
Panel B.4: Recreation — Cajun and Creole establishment indicator				
TV × Hispanic × Cajun and Creole brand	-0.187 (1.630)	-0.187 (1.630)	-0.187 (1.630)	-0.187 (1.631)
County log(income)	Yes	Yes	Yes	Yes
County % Hispanic	No	Yes	Yes	Yes
County log(pop.)	No	No	Yes	Yes
County FE	No	No	No	Yes
NAICS code FE	No	No	No	Yes

Notes: Standard errors are clustered at the county level.

