L. Guillermo Woo-Mora

PhD Candidate in Economics, Paris School of Economics & EHESS
48 Boulevard Jourdan, 75014 Paris, France
guillermo.woo-mora@psemail.eu · woomora.github.io

September 15, 2026

Emily Breza (Harvard University)
Davide Viviano (Harvard University)
Arun Chandrasekhar (Stanford University)
National Bureau of Economic Research
Cambridge, MA

Re: Application to the Workshop on Development Economics and Networks for Early-Career Researchers (December 5, 2026)

Dear Professors Breza, Viviano, and Chandrasekhar,

I am writing to apply to the NBER Workshop on Development Economics and Networks for Early-Career Researchers. I am a PhD candidate in Economics at the Paris School of Economics (PSE) and EHESS, on the job market this year.

I would like to present my job market paper, *"On the Other Side of the Creek: Historical Exclusion, Neighborhood Orientation, and Human Capital Gaps."* Using a colonial natural experiment of neighborhood formation in Guadalajara, Mexico, I show that children on the historically excluded side of a former creek are 3.5 percentage points less likely to attend school, even though measured school supply and quality no longer differ at the boundary. Survey evidence points to a stronger orientation of social relations toward the immediate neighborhood on that side, and an incentivized experiment shows that making the neighborhood salient reduces parents' demand for tutoring by about one fifth there.

Enclosed please find my CV, a brief statement of motivation and research interests, and my job market paper. Should I be shortlisted, my advisor can provide a recommendation:

Suanna Oh, Paris School of Economics · suanna.oh@psemail.eu

Thank you for your consideration. I would be delighted to take part in the workshop.

Sincerely,

L. Guillermo Woo-Mora

Statement of Motivation and Research Interests

L. Guillermo Woo-Mora

I am a job market candidate at the Paris School of Economics. I am applying to the Workshop on Development Economics and Networks for Early-Career Researchers because I want to deepen my understanding of, and gain expertise in, the empirical networks literature, which has become central to the questions I now study.

My job market paper asks whether a neighborhood's social environment can reduce households' investment in education. I study a historical natural experiment of neighborhood formation in Guadalajara, Mexico: at the city's founding in 1542, the conquered Indigenous population was resettled east of a creek, apart from Spaniards and their Indigenous allies on the western bank. Using a spatial regression discontinuity design, I find that nearly five centuries later children on the historically excluded side are 3.5 percentage points less likely to attend school, although measured school supply and quality no longer differ at the boundary. Historical records and present-day surveys point instead to a stronger *neighborhood orientation* of social relations on that side: social interaction, trust, and mutual assistance are more concentrated among people who live nearby. A simple framework predicts that, when educational investment trades off with locally valuable social ties or local schooling norms, making the neighborhood salient should lower educational demand. An incentivized experiment confirms this: randomly making the neighborhood salient before parents choose between mathematics tutoring for their child and a gift card reduces tutoring demand by about one fifth on the historically excluded side. Beyond Guadalajara, places whose Facebook friendship networks are more localized also display lower schooling, especially in more segregated settings.

The paper started as an attempt to understand persistence mechanisms in the context of neighborhood effects in a developing country. As the evidence pointed away from the supply side and toward the structure of local social relations, I arrived, somewhat unexpectedly, at the intersection of development and networks. This is also where the paper's open questions lie. My experiment identifies a demand-side response, but it cannot separate the value of local network ties (help, time, mutual insurance that a household risks losing by investing in education) from conformity to local schooling norms.

In the future, I am to distinguishing these channels. Thus I need measuring who is connected to whom, how ties are formed and maintained, and how behavior propagates through them: exactly the tools and designs the workshop focuses on. Feedback from the organizers and participants would be extremely valuable as I plan to collect network data in the same neighborhoods.

My future research agenda aims to better understand how historical events shape demand-side inequalities in developing countries through social norms and informal networks. In ongoing work with Alex Rivadeneira, for instance, I study how historical debt peonage in Mexico shaped long-run aversion toward formal credit, a setting in which informal lending networks are a natural substitute. Once these mechanisms are better understood, my goal is to design and evaluate policies tailored to closing such demand-side gaps, for instance interventions that complement, rather than compete with, the local networks households rely on.

L. Guillermo Woo-Mora

woomora.github.io guillermo.woo-mora@psemail.eu

September, 2026

Education

PhD in Economics, Paris School of Economics (PSE) & EHESS 2021–2027 (expected)
Dissertation advisors: Suanna Oh, Thierry Verdier, and Ekaterina Zhuravskaya
Visiting Student, Princeton University (host: Ellora Derenoncourt) Fall 2026
Visiting Student, University of Chicago Harris (host: Eduardo Montero) Fall 2024

M2 in Analysis and Policy in Economics, PSE 2020–2021

MSc in Economics, Centro de Investigación y Docencia Económicas (CIDE) 2019–2020

BA in Economics, CIDE 2015–2019

Fields

Primary fields: Development Economics, Political Economy
Secondary fields: Economic History, Applied Economics

References

Suanna Oh (PSE)
✉ suanna.oh@psemail.eu

Thierry Verdier (PSE)
✉ thierry.verdier@psemail.eu

Ekaterina Zhuravskaya (PSE)
✉ ekaterina.zhuravskaya@psemail.eu

Eduardo Montero (University of Chicago)
✉ emontero@uchicago.edu

Fellowships & Awards

2026–2027 ACES Dissertation Fellowship, Association for Comparative Economic Studies
2026 Dissertation Fellowship, Economic History Association
2026 BeNA Innovative Research Award, Berlin Network of Labor Market Research (with D. Kamel)
2025 Víctor L. Urquidi Prize, First Place (with L. G. Woo-Gómez)
2024 Premio Banamex de Economía, Honorable Mention (with M. Montoya-Aguirre and F. Daverio-Occhini)
2024 Víctor L. Urquidi Prize, Honorable Mention (with M. Montoya-Aguirre and E. Chávez)
2022 Sir Alec Cairncross Prize, Runner-Up (with L. Baldomero-Quintana and E. de la Rosa-Ramos)
2022 Manuel Espinosa Yglesias Research Prize, First Place
2020 Mexico's National Prize of Public Finance, First Place (with M. Montoya-Aguirre and E. Ortiz-Juárez)
2019 Premio Banamex de Economía, Undergraduate Thesis, First Place

Research

Job Market Paper

"On the Other Side of the Creek: Historical Exclusion, Neighborhood Orientation, and Human Capital Gaps."

Can a neighborhood's social environment reduce households' educational investment? I study a historical natural experiment of neighborhood formation in Guadalajara, Mexico. At the city's founding in 1542, the conquered Indigenous population was resettled east of a creek, apart from Spaniards and their Indigenous allies on the western bank. Nearly five centuries later, children on the historically excluded side are 3.5 percentage points less likely to attend school, although measured school supply

and quality no longer differ at the boundary. Historical records show that educational institutions were concentrated on the western bank, while naming records and present-day surveys show a stronger orientation toward the immediate neighborhood on the historically excluded side. A simple framework predicts that when educational investment trades off with locally valuable social relations or local schooling norms, making the neighborhood salient should reduce educational demand. I test this prediction in an incentivized experiment in which parents choose between mathematics tutoring for their child and a gift card. Randomly making the neighborhood salient before this choice reduces tutoring demand by about one fifth on the historically excluded side. The results show that educational differences can persist after public provision converges and identify neighborhood salience as an active demand-side margin.

Publications

"Unveiling the Cosmic Race: Skin Tone and Intergenerational Economic Disparities in Latin America and the Caribbean," *Journal of Development Economics*, 2026, 179, 103594.

This paper examines how skin tone shapes intergenerational economic disparities across 25 Latin American and Caribbean countries. First, it analyzes the distribution of ethnoracial identities and skin tone, confirming Mestizo predominance and revealing how broad categories obscure substantial phenotypic diversity. Second, it documents non-linear gaps in income and education, with darker skin tones consistently linked to economic disadvantages. A variance decomposition shows that skin tone explains significant within-group variation, offering explanatory power beyond self-reported ethnoracial categories. Finally, using mothers' education as a benchmark, the study provides novel cross-country evidence on skin tone gaps in absolute educational intergenerational mobility, revealing barriers to upward mobility for darker-skinned individuals. Robustness checks with machine-assessed skin tone data from Mexico, incorporating additional parental and contextual controls, confirm these disparities. These findings underscore the need to account for phenotypic variation when studying economic inequality in Latin America.

"Populism's original sin: Short-term populist penalties and uncertainty traps," *European Economic Review*, 2025, 172, 104917.

This paper investigates the immediate economic impacts of populist policies. In 2018, the Mexican president-elect held an unofficial referendum with less than 1% turnout to halt the construction of Mexico City's New International Airport. I show the policy is plausibly a natural experiment with nationwide impacts, preceding other macroeconomic shocks and populist policies dismantling economic and institutional checks. Using the synthetic control method, I find a 3.3% to 4.5% GDP reduction one year post-cancellation. Consistent with the uncertainty trap framework, this decline is due to heightened economic uncertainty and a significant drop in investment, reinforcing each other and leading to a prolonged economic downturn. The results highlight how populist policies can impose immediate economic costs by creating uncertainty traps, even before implementing institutional changes.

"Infrastructures of race? Colonial indigenous segregation and contemporary land values," with Luis Baldomero-Quintana and Enrique de la Rosa-Ramos. *Regional Science and Urban Economics*, 2025, 110, 104065.

We investigate the persistent impact of a colonial segregation policy on land values in modern Mexico City. During colonial times, Indigenous communities were confined, with varying degrees of success, to settlements known as pueblos de indios. Using historical records, we exploit quasi-random variation due to the pueblos' catchment areas and use a Regression Discontinuity Design to estimate the causal effects of pueblos on land prices. We find a 5% land value penalty for areas affected by the colonial policy. The penalty is exacerbated for the pueblos formerly inhabited exclusively by Indigenous populations. Historical evidence and novel digitized maps reveal that these land value penalties have been driven over the past two centuries by low public goods provision, negative economic expectations, and the historical sorting of working-class individuals who built small housing structures that are second-nature factors. Moreover, in contemporary data, we observe discontinuities in housing overcrowding and public goods quality within the pueblos' catchment areas. Our results underscore the repercussions of colonial policies on contemporary spatial equilibria, clarifying the mechanisms driving historical persistence and offering implications for urban policies.

Working papers

"Moral Force, Contagious Force: Partisanship, Leadership, and Public Health Compliance," with María Montoya-Aguirre and Federico Daverio-Occhini. July 2026.

Does partisan behavior during crises reflect fixed dispositions or malleable responses to elite cues? The two accounts predict similar partisan gaps when leaders reinforce supporters' prior beliefs, but diverge when the same leader changes course. Mexico provides such a reversal: President López Obrador publicly undermined his government's March 2020 distancing campaign, then endorsed vaccination on national television one year later. Combining high-frequency mobility data, electoral returns, and COVID-19 records, we estimate the partisan mobility gap using a difference-in-differences event study, complemented by a recentered lottery instrument for local Morena presence. The dynamics are more consistent with responsiveness to changing cues. Areas with stronger Morena support reduced out-of-home mobility by less after the distancing announcement; the pattern appears across electoral sections, municipalities, and local labor markets, and narrows as local infections rise. Vaccination later increased more in the same areas, with no comparable pattern in routine immunization. Higher-Morena areas also recorded more early cases and deaths, and a dynamic two-group SIRD model attributes roughly one quarter of these conditional gaps to differential mobility. The reversal is difficult to reconcile with fixed dispositions alone and more consistent with responsiveness to changing leader cues.

"Skin Tone Penalties: Quasi-Experimental Evidence on Colorism in Football," with Donia Kamel. May 2026.

We provide causal evidence of skin tone discrimination using professional football (soccer) as a natural laboratory. Leveraging a computer-vision measure of skin tone and quasi-random variation in shot outcomes near the goal frame, we implement a Difference-in-Discontinuities design comparing narrowly scored goals to narrowly missed attempts. We find that Light-skinned players receive significantly larger boosts in post-match ratings than Tan- and Dark-skinned peers for identical actions. These disparities appear in both algorithmic and human-assigned evaluations and are concentrated in the subjective component of ratings. Season-level analyses reveal that biased evaluations translate into lower market valuations for darker-skinned players, despite equivalent performance. Evaluative bias, rather than differential treatment in contracts, emerges as a key driver of economic inequality in this high-information labor market. Our findings show how skin color discrimination can persist even in environments with transparent outcomes and extensive performance data.

"Embracing modernity pays: Cadastre modernization effects on local property tax collection," with María Montoya-Aguirre and Emmanuel Chávez. September 2024.

This paper investigates the impact of the Mexican cadastre modernization program on local property tax revenue. We evaluate a comprehensive modernization initiative, which began in 2010 and included guidelines, technical assistance, and subsidies for local cadastres. Using panel data from 2000 to 2019 and a Difference-in-Differences approach, we find that municipalities adopting the program saw a 10% increase in property tax revenue within five years, rising to 30% after eight years. This effect varies significantly with local state capacity and the coordinating institution implementing the program. The program significantly enhances the state of local cadastres over the long-term. Municipalities that adopt the program are more likely to invest in staff training, implement comprehensive cadastral management systems, and digitize their registries. Our results highlight the importance of local capacities and coordination in realizing the benefits of decentralization and modernization policies.

"Lifting Each Other Up: How Cooperative Firms Foster Local Development," with L. Guillermo Woo-Gómez. October 2025.

Do cooperative firms foster local economic development? This paper examines Mexico, assembling a new georeferenced panel dataset at the 10 × 10 km grid-cell level for 2015–2023. Across multiple research designs, we consistently find that cooperative presence boosts local development, measured by nightlight intensity. An event-study design shows persistent extensive margin gains of 6–8% within four years of entry. To address selection concerns, we implement a shift–share IV strategy based on endogenous lagged-cooperative presence and plausibly exogenous sectoral growth shocks; the results imply that one additional cooperative increases nightlight density by about 2.7%. Complementary census evidence links cooperatives to higher schooling, home asset ownership, and decreases migration pressures. We also document stronger effects in Indigenous communities, where cooperatives

have longer survival spells. Aggregate patterns suggest that cooperatives increase agricultural specialization but with clear productivity gains, while redistributive effects raise local disposable income and decrease poverty. These findings provide novel causal evidence that cooperatives can function as engines of inclusive local development.

"New Echoes of Deep Mexico? Indigenous Identity and Its Enduring Imprints," with Pedro Torres-López. *HKS Project on Indigenous Governance and Development*, Working Paper 26-009, September 2026. Chapter for the second volume of the book *Roots of Underdevelopment: A New Economic (and Political) History of Latin America and the Caribbean*, edited by Felipe Valencia-Caicedo.

This chapter revisits Guillermo Bonfil Batalla's 'México Profundo' (1987) through the lens of contemporary empirical historical economics. In that work, Bonfil Batalla argued that beneath a Westernized facade of modern Mexico lies a "Deep Mexico": a living but denied Indigenous civilization that continues to shape the country's culture, identity, and social life. The first part of the chapter compiles historical census tabulations and harmonized microdata to document persistent patterns of Indigenous identity, inequality, and historical continuity. The historical process of Indigenous identification is multidimensional and endogenous, shaped by language, self-identification, phenotype, and geography. Despite improvements in basic services, Indigenous populations continue to face structural gaps in education and income. Beyond explicit identifiers, the second part of the chapter examines three channels through which 'México Profundo' continues to influence economic life. First, skin tone remains a consistent predictor of disadvantage, reflecting racialized inequality. Second, historical Indigenous presence structures spatial inequality and human-capital outcomes. Third, Indigenous institutions—both formal and informal—continue to shape development trajectories, depending on their interaction with markets and the state. Together, these patterns underscore that 'México Profundo' continues to illuminate key mechanisms linking Indigenous identity to economic and social outcomes in modern Mexico.

Work in progress

"From Indenture to Aversion: Debt Peonage and Long-Run Credit Demand," with Alex Rivadeneira.

Conferences and Presentations

2026: *Race, Ethnicity, and Policies in Latin America and the Caribbean*, IDB, Washington; LSE; PSE; *CEPR Development Economics Annual Symposium 2026-1*, Trinity College Dublin, Ireland; LISER; BeNA Labor Economics Workshop; Linnaeus Inequality Summit; World Inequality Conference; BSE Summer Forum; SIOE; EEA-ESEM; ACES Summer School; Brown*; Princeton*.

2025: *ASSA Annual Meeting*; *Workshop in Economic History*, ENS Lyon; *The Rural World in Latin America and the Caribbean – Workshop*, CAF, Buenos Aires; *LACEA-EHN Workshop on Historical Development*, Universidad del Pacífico, Peru; PSE; *Workshop on Social Economy for Young Economists*, University of Bologna; UNU-WIDER; *Diversity and Human Capital Workshop*, University of Exeter; *Narratives and Culture in Historical Perspective*, LSE; Banco de México; ITAM; PSE; PSL.

2024: CIDE; *PhD - Economics Virtual Seminar*; *Virtual Workshop in Historical Political Economy*; *Equality of Opportunity and Intergenerational Mobility: A Global Perspective*, University of Bari; *Workshop on Cultural Transmission and Persistence*, Queen's University Belfast; *Advances with Field Experiments Conference*, LSE; *Text-as-Data*, Monash-Warwick-Zurich-CEPR; *SobreMéxico*, U. Iberoamericana CDMX; Experimental Lunch seminar, University of Chicago; *ECINEQ-LAC*, Colegio de México.

2023: *Lewis Lab Graduate Student Workshop*, Arthur Lewis Lab; *LACEA-EHN Workshop on Historical Development*, CAF, Montevideo; *LACEA Inequality & Poverty*, CAF, Montevideo; *Urban Economics Association 12th European Meeting*, Bocconi; *Neighborhoods and local interactions*, PSE; *REUS Workshop*, CEU Vienna; *LAUrban Meeting*, FGV.

2022: *Urban Economics Association 11th European Meeting*, LSE; *Applied Young Economist Webinar*, Monash Business School and Warwick; *Doctorissimes*, PSE; *CROSS*, APSA Class and Inequality section; *SobreMéxico*, U. Iberoamericana CDMX.

2021: *World Inequality Conference 2021*, PSE; *VPDE PhD Workshop*. Colegio Carlo Alberto. Turin, Italy; *The Public and Democracy in the Americas 2021*, LAPOP Lab and ITAM; *DEMOSAL*, Colegio de México.

2020: *RIDGE Workshop on Economic History*, RIDGE; *RIDGE Workshop on Inequality and Poverty*, RIDGE and LACEA NIP; *PhD Student Workshop in Urban Economics*, Urban Economics Association.

* Scheduled presentations.

Research Grants

PhD Research Funds – PSE Stone Center on Global Wealth Dynamics (2025).

Mondialisation, développement et environnement – CEPREMAP (2025).

Inequality over the life course – UNU-WIDER (2024 – 2025).

The Rural World in Latin America and the Caribbean – CAF (with L. G. Woo-Gómez, 2024 – 2025).

Descentralización y Capacidades Subnacionales en América Latina y el Caribe – CAF (with M. Montoya-Aguirre and E. Chávez, 2023 – 2024).

Arthur H. Cole Grant in Aid – Economic History Association (with L. Baldomero-Quintana and E. de la Rosa-Ramos, 2021).

Dissertation Grant – American Real Estate and Urban Economics Association (2021).

Teaching Experience

Adjunct Lecturer (ATER, teaching load 60 hours) - Université Paris Sciences et Lettres (PSL) (2024-2026):
- Microeconomics II
- CORE Economics Macro

Adjunct Lecturer - PSL (2022-2023):
- CORE Economics Micro
- CORE Economics Macro

Skills

Software: R, Python, Stata, QGIS.

Languages: Spanish (native speaker); English (fluent); French (intermediate).

On the Other Side of the Creek:

Historical Exclusion, Neighborhood Orientation, and Human Capital Gaps*

L. Guillermo Woo-Mora†

Job Market Paper

Version: September 14, 2026

Click here for the most recent version.

Abstract

Can a neighborhood's social environment reduce households' educational investment? I study a historical natural experiment of neighborhood formation in Guadalajara, Mexico. At the city's founding in 1542, the conquered Indigenous population was resettled east of a creek, apart from Spaniards and their Indigenous allies on the western bank. Nearly five centuries later, children on the historically excluded side are 3.5 percentage points less likely to attend school, although measured school supply and quality no longer differ at the boundary. Historical records show that educational institutions were concentrated on the western bank, while naming records and present-day surveys show a stronger orientation toward the immediate neighborhood on the historically excluded side. A simple framework predicts that when educational investment trades off with locally valuable social relations or local schooling norms, making the neighborhood salient should reduce educational demand. I test this prediction in an incentivized experiment in which parents choose between mathematics tutoring for their child and a gift card. Randomly making the neighborhood salient before this choice reduces tutoring demand by about one fifth on the historically excluded side. The results show that educational differences can persist after public provision converges and identify neighborhood salience as an active demand-side margin.

Keywords: Neighborhood Effects, Educational Investment, Historical Persistence, Social Capital.
JEL: *I24*, *R23*, *Z13*, *N36*, *O15*.

*I am deeply grateful to my PhD advisors, Suanna Oh, Thierry Verdier, and Ekaterina Zhuravskaya, for their superb guidance throughout this journey. I thank Livia Alfonsi, Milena Almagro, Siwan Anderson, Eva O. Arceo-Gómez, Luz Marina Arias, Ragui A. Assaad, Samuel Bazzi, Celina Becerra Jiménez, Marie Beigelman, Hernán Bejarano, Guillaume Blanc, Dan Bogart, Davide Cantoni, Felipe Castro Gutiérrez, Denis Cogneau, Ellora Derenouncourt, Steven Durlauf, James Fenske, Paola Giuliano, Laurent Gobillon, Camille Hémet, Luka Henkel, Soeren Henn, Richard Hornbeck, Antoine Hubert de Fraisse, Erin M. Kelley, Rachel Kranton, Sylvie Lambert, Jeffrey Lin, John List, Sara Lowes, Karen Macours, David Margolis, Stelios Michalopoulos, Mushfiq Mobarak, Eduardo Montero, María Montoya-Aguirre, Héctor Paredes Castro, Torsten Persson, Thomas Piketty, Diego Ramos-Toro, Alex Rivadeneira, James Robinson, Guilherme Rodrigues Lambais, Irvin Rojas, Esteban Rossi-Hansberg, Michaela Sviatchi, Desislava Tartova, Felipe Valencia-Caicedo, Randall Walsh, Geoffrey Wodtke, Liam Wren-Lewis, David Yang, and Yves Zenou for insightful conversations and feedback. I also thank participants at several conferences and seminars for their valuable comments. This work benefited from the support of EHESS and the ANR (ANR-17-EURE-0001), as well as financial support from ACES, EHA, UNU-WIDER, CEMPREMAP, the PSE Stone Center, and AREUEA. All remaining errors are my own.

†PhD Candidate. Paris School of Economics, France. guillermo.woo-mora@psemail.eu.

1 Introduction

The place where children grow up has far-reaching consequences over the life cycle. Children raised in disadvantaged neighborhoods attain less education, earn less, and experience lower social mobility, making neighborhoods an important source of inequality (Durlauf, 1996; Chyn and Katz, 2021; Katz, 2026). Neighborhood effects have traditionally been traced to material features of place, such as schools, labor-market opportunities, and housing.

But places contain people as well as infrastructure. Recent evidence points to a neighborhood's social environment as a distinct input into children's outcomes. Consistent with arguments from sociology, changes in childhood communities track changes in mobility, and public-housing redevelopment can improve outcomes by connecting residents to surrounding social networks (Sampson et al., 2002; Chetty et al., 2026b; Jackson, 2026). Neighborhoods therefore bundle material conditions with a social environment.

This paper asks whether a neighborhood's social environment can reduce households' educational investment. I focus on the spatial orientation of social relations: the extent to which social interactions, trust, and mutual assistance are concentrated among people who live nearby rather than extending beyond the immediate neighborhood. I refer to this dimension as *neighborhood orientation*.

The effect of neighborhood orientation on educational investment is not obvious. Local ties can provide help, time, and other forms of support. But educational investment can also pull households away from those ties or from the schooling choices of people around them. It is an open empirical question whether this trade-off exists and whether it matters in practice.

Answering this question is challenging, since neighborhoods bundle their social environment with the material conditions of place. Most experimental and quasi-experimental evidence changes both margins at once: families move to places with different schools, public goods, and opportunities, but also a different social pool.

The city of Guadalajara, Mexico, provides a setting in which historical differences in educational provision have disappeared while a longstanding neighborhood boundary remains socially salient. A colonial boundary divided the city across a creek, later paved into an avenue. The two sides differed sharply in educational provision historically, but today share the same labor market and schooling system, with no systematic difference in measured school supply or quality at the boundary. Using a spatial regression discontinuity design (RDD) around the former creek, I find that residents attain fewer years of schooling and children on the historically excluded side are 3.5 percentage points less likely to attend school.

I then turn to the demand side, asking whether educational choices respond to the neighborhood environment while public provision remains fixed. Historical records and present-day surveys point to a stronger neighborhood orientation on the historically excluded side, while aspirations, expectations, and perceived returns to schooling are similar across the boundary. A simple

framework links neighborhood orientation to educational demand, and I test its salience prediction in an incentivized choice experiment. Making the neighborhood salient before an educational choice reduces tutoring demand by about one fifth on the historically excluded side. The paper provides novel evidence on a demand-side margin of educational investment in a historically excluded neighborhood where social relations remain more locally oriented.

To study whether the social environment matters after differences in public provision disappear, I exploit a historical natural experiment of neighborhood formation. At Guadalajara's foundation in 1542, Spaniards and their Indigenous allies settled west of a creek, while the conquered local Indigenous population was resettled east. The boundary generated a sharp difference in colonial status within a common urban setting. Because these settlements formed the seed of the city, the two populations entered the same new urban setting, with similar locational fundamentals and no pre-existing neighborhoods on either side. The difference in status did not take the form of a separate legal regime, and for most of the colonial period both banks fell under the same municipal authority.

The colonial distinction later disappeared, and educational provision converged across the boundary. In the early twentieth century, the creek was piped underground and replaced by an avenue, the Calzada Independencia. Today, both sides share the same jurisdiction and face no school catchment boundary or physical barrier to movement, and the historical divide is no longer an ethnoracial one. The boundary nonetheless survives in ordinary speech, where residents locate themselves and others with the expression *on the other side of the Calzada*, which refers to the eastern bank.

I estimate the long-run effect of assignment to the historically excluded side using a spatial RDD around the former creek. Identification requires potential outcomes to vary smoothly across the boundary in the absence of the historical assignment. This assumption is plausible because the boundary predates the city around it, locational fundamentals are similar across banks, and today it coincides with no jurisdiction, school catchment, physical barrier, or ethnoracial divide. All spatial RDD estimates condition on age and sex, so the discontinuities do not reflect differences in basic demographic composition across the two sides. Residential location is chosen, so the design cannot rule out sorting on unobserved determinants of education; I therefore examine housing values, migration, and other present-day differences explicitly.

The historical boundary still predicts human capital nearly five centuries later. Children on the historically excluded side are 3.5 percentage points less likely to attend school (0.09 standard deviations), while residents attain about 0.8 fewer years of schooling (0.17 standard deviations). The attendance gap is especially informative because it reflects a decision being made today, after measured school provision has converged across the boundary. Administrative records show no current discontinuity in school supply or quality, and the gap is not accompanied by a discontinuity in crime exposure or explained by housing values or household income. The gap is also largest where Analco borders Mexicaltzingo, an Indigenous settlement whose residents did

not share Analco's excluded colonial status. Because both sides of this segment were historically Indigenous, ethnicity alone is difficult to reconcile with the spatial pattern. The educational divide is already visible in the 1930 census and remains visible in later waves.

The provision of public goods differed historically across the boundary, while it is no longer visible in contemporary school data. I georeference the indices of four city maps between 1800 and 1908 and locate each named institution relative to the creek. The eastern bank accounted for between a third and two-fifths of the population in the available enumerations but for less than fifteen percent of the institutions named in every map. No educational establishment appears in the eastern map indices until 1908. The maps do not measure the full supply of schooling, but they show that the city's named educational institutions were concentrated on the western bank.

Relational differences across the boundary were already visible by the end of the colonial period and remain visible in different forms today. Men east of the creek were more likely to bear the name of their parish's patron saint in the 1821 census, a difference that separates Analco from Mexicaltzingo as well as from the Spanish west. Historical records show that given names and paternal surnames remained distinctive across the boundary into the twentieth century. Using present-day surveys, I show that residents east of the boundary trust their neighbors substantially more, trust formal institutions less, and are more likely to belong to a neighborhood association, even conditional on schooling. These measures point to a stronger orientation toward the immediate neighborhood on the historically excluded side.

I first develop a simple framework linking neighborhood orientation to educational demand. Educational investment raises market returns but can trade off with the neighborhood through two margins. The first is a time-investment margin: time devoted to education reduces time spent with neighbors, which is needed to maintain local ties and obtain their benefits. The second is a social-conformity margin: educational choices that depart from local schooling behavior carry a psychological cost. When the resulting neighborhood pull is positive, making the neighborhood salient reduces educational demand.

I test this prediction in a pre-registered incentivized choice experiment. Parents make an incentivized choice between an hour of mathematics tutoring for their child and a gift card. As part of the survey, participants complete a mapping exercise in which they use a tablet to draw the boundaries of what they consider their neighborhood and explain why they define it that way. I randomize whether they complete this exercise immediately before or after the educational choice. The manipulation provides no information and changes no incentives: it varies only whether participants have just actively delineated and described their neighborhood when they choose. On the east bank, mapping the neighborhood first makes parents 12 percentage points less likely to choose tutoring, about a fifth of the control mean. Willingness to pay falls from MXN 176 to MXN 102, about two fifths. On the west bank, both estimates are positive and imprecise.

The first experiment makes the neighborhood salient, but not the historical meaning of the

boundary. A second pre-registered exercise therefore makes the Calzada divide itself salient before respondents explain why students in their neighborhood may not continue their education. Within Analco, the prime reduces references to money and work constraints by 15 percentage points. It also shifts the relative balance of explanations away from economic constraints and toward neighborhood-based explanations.

Complementary survey modules provide descriptive evidence on what may underlie this response to neighborhood salience. Analco residents receive more help from neighbors, spend more time with them, and direct more giving toward a stranger from their own neighborhood. By contrast, parents on the two sides report similar educational aspirations, perceived monetary returns to schooling, and explicit work-versus-study norms. Analco respondents also require an approximately 18% higher salary before they consider university worthwhile, consistent with an additional nonpecuniary hurdle to educational investment. These comparisons help characterize what differs across the two sides, but they do not identify which feature of the neighborhood drives the experimental response.

A dynamic extension provides an account of why the educational and relational differences can persist after public provision converges. Unequal historical provision can initially generate an educational gap, after which education and neighborhood orientation can become mutually reinforcing. With sufficiently persistent intergenerational transmission of neighborhood orientation, as in models of cultural transmission (Bisin and Verdier, 2000, 2001, 2026), and slow residential turnover, these differences can survive the convergence in public provision. Long-term residence is more common east of the boundary, consistent with the low-turnover condition in the model. In 1992, a series of explosions destroyed part of Analco, leading to substantial rebuilding and residential turnover. The affected blocks initially attracted more educated residents and had higher child attendance conditional on parental schooling, but both differences faded in later censuses. Because the episode bundled displacement, reconstruction, and changes in housing, I interpret it as consistent with the framework rather than as an identified test of its dynamics.

The pattern documented in Guadalajara is not unique to that setting. Using the share of Facebook friendships that remain within a short radius as a measure of the localization of social ties, I find that more locally concentrated ties are associated with lower schooling across different geographic levels worldwide. The individual-level evidence from the World Values Survey shows the same pattern more directly. Across 41 countries, individuals with a higher neighborhood-trust premium attain less schooling, are less likely to surpass their parents' education, and are more likely to leave school before completing upper secondary. These correlations are concentrated in highly segregated settings, where the schooling gradient is larger than in less segregated settings. These comparisons are descriptive and may reflect selection, omitted variables, or other differences across places. Their value is that the same broad relationship that appears in Guadalajara also appears across very different institutional and geographic settings: more locally oriented social relations are associated with lower human capital, especially where residential segregation is

greater.

This paper makes two primary contributions. First, it identifies a demand-side margin of neighborhood effects by experimentally varying neighborhood salience while holding residential location and material conditions fixed. In the historically excluded neighborhood, where social relations are more locally oriented, salience reduces educational demand. Second, it studies persistence after material conditions converge. In Guadalajara, historical differences in educational provision disappeared, while educational and relational differences remain visible across the former boundary. A simple framework links these patterns through feedback between education and neighborhood orientation.

Related literature. The first contribution relates to work on neighborhood effects (Durlauf, 2004; Topa and Zenou, 2015; Chyn and Katz, 2021; Katz, 2026). A large literature shows that where children grow up shapes their adult outcomes (Kling et al., 2007; Ludwig et al., 2013; Chetty et al., 2016; Chyn, 2018; Nakamura et al., 2022; Bergman et al., 2024), with recent work emphasizing the social environment as an important part of neighborhood opportunity (Jackson, 2026; Chetty et al., 2022, 2026a). Much of the causal evidence comes from moving families, improving access to higher-opportunity neighborhoods, or otherwise changing the environments to which families are exposed. This paper instead holds the family and residential location fixed and varies neighborhood salience at the moment of an educational choice. This margin differs from gaining access to a better-connected community (Chetty et al., 2026a) or losing local ties through displacement (Rojas-Ampuero and Carrera, 2026). It is related to evidence that peers and social image affect educational choices (Fryer and Torelli, 2010; Bursztyn and Jensen, 2015; Bursztyn et al., 2019), but here the salient social reference is the neighborhood and the decision is made by parents rather than students. Because participants delineate their own neighborhoods, the design also does not impose the administrative definitions commonly used to measure neighborhood effects (Aliprantis, 2023; McCartan et al., 2024).

The second contribution relates to work on historical development and persistence (Nunn, 2014, 2020; Voth, 2021). Long-run differences across places have been traced to institutions (Acemoglu et al., 2001; Dell, 2010), geography (Davis and Weinstein, 2002; Bleakley and Lin, 2012), early human capital (Valencia-Caicedo, 2019), and culturally transmitted traits (Nunn and Wantchekon, 2011; Nunn, 2012; Michalopoulos and Xue, 2021; Lowes, 2023a; Michalopoulos et al., 2025). A growing urban literature further shows that historical shocks can leave durable spatial footprints within cities (Ambrus et al., 2020; Heblich et al., 2021; Hanlon and Heblich, 2022; Lin and Rauch, 2022; Baldomero-Quintana et al., 2025), often through persistent features of place such as the built environment, infrastructure, land use, and spatial sorting. This paper instead studies persistence in the behavior and social relations of people living in historically separated neighborhoods, connecting the urban persistence literature to work on the long-run effects of historical institutions on trust, social preferences, and other culturally transmitted traits (Grosfeld and Zhuravskaya,

2015; Lowes and Montero, 2021; Lowes, 2023b; Ramos-Toro, 2023). Guadalajara makes it possible to study persistence after the original institutional and educational asymmetries have disappeared. The framework rationalizes this persistence through an intergenerational feedback between education and neighborhood orientation. The historical evidence is consistent with this process, while the experiment shows that the social environment remains behaviorally relevant today.

The paper also relates to work on social capital. Social capital encompasses distinct dimensions of networks, trust, and participation, and these dimensions need not move together (Durante et al., 2025). Broad measures of social capital are often positively associated with development (Knack and Keefer, 1997; Goldin and Katz, 1999; Guiso et al., 2004; Algan and Cahuc, 2014), while close or cohesive social ties can also impose costs (Portes, 1998; Tabellini, 2008; Enke, 2019; Cappelen et al., 2025; Le Rossignol and Lowes, 2026). Rather than measuring the amount of social capital, I focus on its spatial organization: whether social relations are concentrated in the immediate neighborhood rather than extending beyond it. The evidence from Guadalajara and the cross-setting exercises shows that this local orientation is associated with lower schooling, while the experiment establishes that making the neighborhood salient can reduce educational demand in the setting where neighborhood orientation is stronger.

Finally, I contribute to the literature on identity economics (Akerlof and Kranton, 2000, 2002; Shayo, 2020; Atkin et al., 2021; Oh, 2023). Most of this literature focuses on social categories such as race, ethnicity, or caste, with less attention to geographic space. My experimental evidence shows that the neighborhood itself can operate as a salient social category and affect economic decisions.

The rest of the paper proceeds as follows. Section 2 describes the historical setting and the colonial boundary. Section 3 estimates its long-run effects on human capital. Section 4 documents the historical convergence in educational provision and the relational differences that remain across the boundary. Section 5 develops the framework and presents the experimental and survey evidence on educational demand, together with suggestive evidence on persistence. Section 6 examines related patterns beyond Guadalajara. Section 7 concludes.

2 Context and Historical Natural Experiment

Setting. Guadalajara, in western Mexico, is a rich, highly schooled municipality by national standards and anchors the country's second-largest metropolitan area.[1] The gap I study is therefore

[1] In 2020, average quarterly household income was 68,308 pesos, 36 percent above the national average and at the 98th percentile among urban municipalities. Residents averaged 10.96 years of schooling, compared with 9.74 nationally, and school attendance beyond primary has remained above the national rate for decades. Figure A.1 in Appendix A locates Guadalajara in the national distribution of each outcome, and Figure A.2 shows that the municipality has exceeded the national attendance rate at ages 12 and above in every census since 1970. Primary-age attendance is now near-universal everywhere, and Guadalajara sits marginally below the national rate on that margin.

not between a poor place and a rich one, but within a wealthy, highly schooled city.

Much of the variation in school attendance in Mexico occurs within municipalities. Among Mexicans aged 12–17, differences between municipalities account for only 3.7 percent of the variance in attendance in 2020, while differences between households within the same municipality account for 78.7 percent.[2] I study this within-municipality variation in Guadalajara, where human capital varies sharply across census blocks as Figure 1b shows. Across the historical boundary, residents share one labor market (Aldeco et al., 2024) and face no school catchment areas or physical barriers to movement, so these features are common to both banks.

Foundation. Spanish authorities founded Guadalajara in the Atemajac Valley in February 1542 as the capital of New Galicia, on land with no existing settlement. Because there was no town to inherit, the same authorities decided where Spaniards and the conquered Indigenous population would live and recorded both assignments in a single administrative act.

Unlike central Mexico, where dense centralized polities were absorbed into the colonial order through co-optation of the Mexica (*Azteca*) state (Knight, 2002; Arias and Dirod, 2014), New Galicia was sparsely populated and politically fragmented. Fragmentation made the conquest slow: with no central authority to co-opt, each group had to be defeated separately, and no surrender bound the rest. Three earlier attempts at founding the city collapsed under Indigenous resistance and water shortages.[3]

That act drew on a Crown-wide institutional architecture of segregation. Beginning in 1538, the Spanish Crown established *Repúblicas de Indios*, a form of corporate and administrative organization for Indigenous populations, typically endowed with their own *cabildos* and communal land rights (Castro Gutiérrez, 2010; Graham, 2013). These institutions were distinct from *Pueblos de Indios*, the geographic settlements in which Indigenous populations lived: a *Pueblo* need not itself constitute a separate *República*. The system sought to separate Indigenous and Spanish populations spatially, but in Guadalajara enforcement was partial (Anderson, 1988). The institutional status of the *Repúblicas de Indios* was abolished after Mexican independence in 1821.

Within Guadalajara, this colonial architecture took a specific spatial form. Spaniard Guadalajara was founded on the western bank of the San Juan de Dios (SJdD) creek. Mendoza's Indigenous allies, Mexicas and Tlaxcaltecas, settled immediately to the south, also on the western bank, in *Mexicaltzingo* ("house of the Mexicas"). One year later, in 1543, Franciscans carried out the

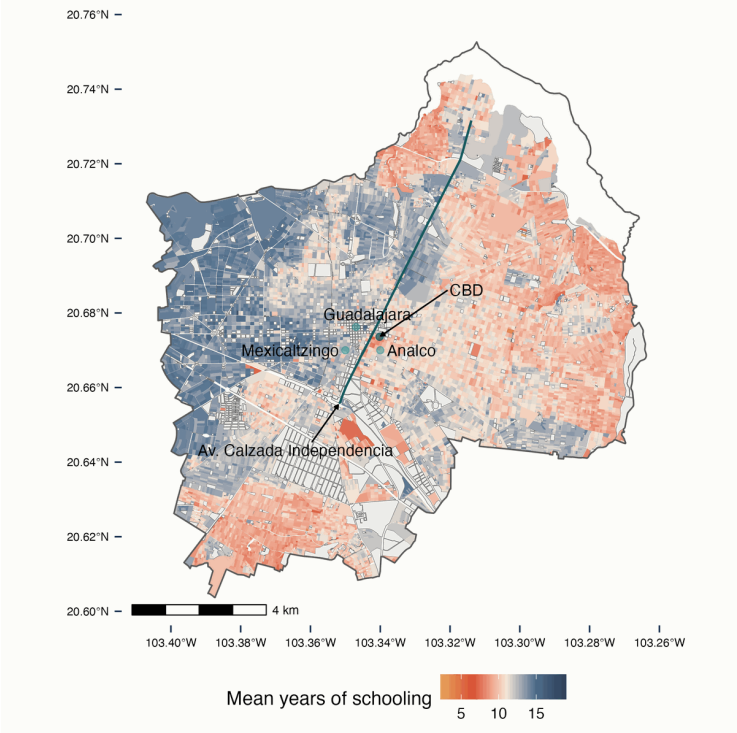
[2] Age, sex, parental schooling, birth order, and sibship size together explain 13.7 percent of the total variance. Because the harmonized census records no geography below the municipality, the household component also contains differences between neighborhoods. Figure A.3 in Appendix A reports the decomposition.

[3] The first settlement (1532, Nochistlán) collapsed after one year; attempts followed in Tonalá (1533) and Tlacotlán (1535). The fourth and final foundation took place in 1542 at the present-day site, south of the Huentitán Canyon, after the *Mixtón* War (1541) and a decisive expedition led by Viceroy Antonio de Mendoza with critical participation of Tlaxcalteca and surrendered Mexica allies.

Figure 1: The foundational settlements in 1542 and schooling in 2020



(a) The three foundational settlements, 1542



(b) Mean years of schooling by census block, 2020

Notes: Panel (a) reconstructs the three foundational settlements from Tanck de Estrada (2005) and IMEPLAN (2016); the thick blue line is the SJdD creek as documented in 17th-century accounts (Lázaro de Arregui, 1946) and 19th-century cartography. Settlement boundaries are illustrative, since jurisdictional delineations between *Pueblos* were not always recorded. Panel (b) maps mean years of schooling by census block in the 2020 *Censo de Población y Vivienda*, shaded from red (lower) to blue (higher), with Calzada Independencia, the avenue that replaced the piped-over creek, overlaid. Blank blocks are those for which INEGI suppresses statistics because of small population counts.

assignment east of the creek, establishing *Analco* ("on the other side of the creek" in Nahuatl) as a *Pueblo de Indios* for the conquered local population (Solís Matías, 1986, 1992). Figure 1a maps the three settlements, all located within roughly one kilometer of each other.

The two Indigenous settlements were not symmetric, and that asymmetry is central to the identification strategy. Mexicaltzingo's settlers participated in the conquest of New Galicia and were recognized as allies, receiving exemption from tribute, communal land grants, and recognition of their leaders as *principales*. They were "*not conquered, but among the conquerors*" (Castro Gutiérrez, 2010) and settled on the same western bank as Spaniard Guadalajara. Analco, the only foundational settlement east of the creek, was populated by conquered local groups and later migrants from the region. For most of the colonial period, neither settlement held the corporate autonomy of a *República de Indios*, and both fell under the city's *ayuntamiento*. The east–west contrast is therefore one of colonial status within a shared legal regime, with conquerors and their Indigenous allies on the west and the conquered population on the east, rather than a simple Spanish-versus-Indigenous divide or a jurisdictional boundary.[4] By 1821, both banks were ethnically mixed but still differed in rank (Appendix A.6).

Colonial exclusion. The historical natural experiment of neighborhood formation isolates the persistence of colonial *exclusion* rather than ethnicity itself. In this setting, the deprivation described in Section 1 took an administrative form: an authority assigned a group to a status that denied it the standing granted to comparable groups within the same jurisdiction (Sen, 2000; Ramos-Toro, 2023). I examine two dimensions associated with that exclusion: the public provision withheld from the group's territory and differences in the organization of local social relations.

A historical natural experiment of neighborhood formation. Five features support the interpretation of this configuration as a credible historical natural experiment in the sense of Cantoni and Yuchtman (2021); Appendix A reports the supporting evidence in detail.

First, settlement locations were unrelated to pre-colonial agglomerations. Locational fundamentals and early agglomerations shape the spatial distribution of economic activity over very long horizons (Davis and Weinstein, 2002; Bleakley and Lin, 2012; Maloney and Valencia-Caicedo, 2016; Lee and Lin, 2018), so a site chosen for such advantages would confound the colonial assignment with the conditions that attracted settlement in the first place. The region had no urban centers prior to Spanish contact, and pre-colonial population estimates (Goldewijk et al., 2017) and colonial-era urban density data (Buringh, 2018) show that it remained sparsely populated relative to central Mexico through the 16th–18th centuries (Figures A.4 and A.5). The three

[4] Reading the contrast through modern ethnic categories would be anachronistic. Alberro (2019) argues that colonial New Spain classified people by the categories of its own time — estate, *nación*, and corporate membership — rather than by race, and that the category *indio* was itself porous. The distinction the sources document between Mexicaltzingo and Analco is one of standing within that corporate order, not of ethnic origin.

foundational settlements were therefore established on previously unurbanized land.

Second, placement was not driven by geographic fundamentals. Altitude, slope, and climate are similar on the two banks, and a regression discontinuity within a 1.5-km buffer of the SJdD creek finds no significant differences in locational characteristics across banks (Figure A.6 and Table A.1).

Third, the three settlements shared the same higher-level political and institutional framework from the city's foundation (Acemoglu et al., 2001, 2005; Dell, 2010; Lowes et al., 2017). They fell under the same colonial authorities centered in Guadalajara and the Audiencia of New Galicia (Lázaro de Arregui, 1946; Solís Matías, 1992), even though Indigenous settlements could differ in their local corporate status and privileges.[5] Historically and today, all three have remained within the same overarching jurisdictional framework, sharing the same court system, fiscal authority, and policing.

Fourth, the foundation predates the industrial, infrastructural, and demographic shocks that reshaped Guadalajara in the 19th and 20th centuries, serving as the city's *seed* rather than as one shock among many (Ahlfeldt et al., 2015; Heblich et al., 2021; Hornbeck et al., 2025).

Fifth, the boundary is not a physical barrier today. The SJdD creek was piped underground in the early 20th century and replaced by Calzada Independencia, a surface-level avenue integrated into the city grid (Appendix A.4). Unlike the railroad tracks and urban freeways studied in the U.S. literature on infrastructural barriers (Ananat, 2011; Chyn et al., 2026; Bagagli, 2025; Valenzuela-Casasempere, 2025), the Calzada connects the two banks rather than separating them, so a discontinuity at the boundary is unlikely to reflect the cost of crossing it.

Together, these features make the colonial assignment a credible source of variation in exclusion, holding the physical, institutional, and market environment common across the two banks.

The boundary today. Although it is no longer physical, the boundary survives as a category in ordinary speech. Residents of Guadalajara locate themselves and others with the expressions *de la Calzada para allá* and *del otro lado de la Calzada*, both of which name the eastern bank (Jaramillo-Molina and Saucedo, 2016; Jaramillo-Molina, 2022). Trujillo Bretón (2018, pp. 310–311) documents that the early twentieth-century press organized the city through a west–east dichotomy bounded by the same creek and assigned disrepute to the eastern side. Reguillo (1998) describes the piped creek as a symbolic mark that still divides the city, and her ethnography of Analco reports residents who locate risk outside the neighborhood, organize exchange through neighbors, and have lived there for thirty years on average.

[5] Around 1800, Analco and Mexicaltzingo are two of four *pueblos de indios* near Guadalajara that do not appear in the *propios y arbitrios* accounts submitted to the Contaduría. Tanck de Estrada (1999) reads that absence as evidence that they fell under the jurisdiction of the city's *ayuntamiento*.

3 Persistent Neighborhood Effects on Educational Outcomes

Nearly five centuries after the colonial assignment, the boundary still divides human capital at a fine spatial scale, despite the disappearance of the institution and the convergence in measured school supply and quality across the two sides. It divides both the stock of schooling residents have accumulated and the schooling decisions their households make today.

3.1 Geolocated census microdata

The main analysis uses restricted-access individual-level census microdata from Mexico's national statistical agency (INEGI).[6] The 2020 *Censo de Población y Vivienda* covers the universe of Guadalajara residents and identifies their census block, the finest spatial unit available in the restricted microdata. This combination of full population coverage and high spatial precision allows me to implement the spatial RDD within meters of the historical boundary.

The census contains limited information on income and labor-market outcomes. To complement it, I use the 2015 *Encuesta Intercensal*, a large-scale survey representative at the municipal level that records detailed income, parental, and household variables. The Intercensal has fewer observations near the boundary than the census, so I use it for municipality-wide comparisons.

3.2 Estimating the discontinuity at the colonial boundary

Specification. To evaluate the long-run effect of the historical division on human capital, I implement a spatial RDD around the SJdD creek, today Calzada Independencia. Location is two-dimensional, so the baseline specification collapses it into the signed orthogonal distance to the boundary, the distance-based approach of Cattaneo et al. (2026):

$$y_i = \alpha + \tau\, \mathbb{1}\{\text{Analco}_i\} + \delta s_i + \zeta\, \mathbb{1}\{\text{Analco}_i\} \cdot s_i + \gamma' X_i + \varepsilon_i. \tag{1}$$

The outcome y_i is the human capital variable of interest for individual i. The treatment indicator $\mathbb{1}\{\text{Analco}_i\}$ equals one for residents east of the boundary: historically the Analco side, today east of Calzada Independencia. The running variable s_i is the orthogonal distance in meters to the boundary, signed so that positive values fall on the eastern side. The interaction lets the outcome's gradient in distance differ across banks, so that τ is identified from a local polynomial fit with side-specific slopes. The control vector X_i includes age and sex. Parental education and household size are added only in the specifications labeled as such. Household income is available only in the Intercensal Survey and therefore enters only those specifications in Table 1; the census-based boundary estimates are not income-adjusted. I select bandwidths and compute robust inference following Calonico et al. (2014), and include boundary-segment fixed effects so that comparisons occur within the same segment, in the spirit of Dell (2010).

[6] These data are not publicly available and can be accessed only through INEGI's *Laboratorio de Microdatos*, which provides researchers access to anonymized census records on its secure servers.

What τ estimates. Treatment changes along a line, not at a point, so the design defines an effect at every boundary point, $\tau(b)$ (Keele and Titiunik, 2015). Equation (1) can be interpreted as a weighted average of effects along that curve. I also report bivariate estimates following Cattaneo et al. (2025) that use both coordinates and trace $\tau(b)$ along the creek.

Identifying assumptions. Identification rests on continuity: average potential outcomes are continuous in location at each boundary point. This assumption is more demanding in geographic designs because residential location, unlike the running variable in many RDD applications, is chosen. Lee and Lemieux (2010) note that "economic agents have quite precise control over the location to place a house or where to live," and Keele et al. (2017) emphasize that residents may have sorted around a boundary for decades.

Households in Guadalajara do choose which side of the avenue to live on. I therefore use density tests (McCrary, 2008; Cattaneo et al., 2018) to examine the distribution of residential population around the historical boundary (Figure B.1). Residential density falls toward the boundary from both sides, reflecting the location of Guadalajara's historic commercial and administrative center, where residential population is sparse. At the cutoff, estimated population density is higher on the eastern side, although the difference narrows to about four percent by 500 meters. This pattern is consistent with the land-use transition at the boundary: the western edge contains the commercial and administrative core, while residential land use begins immediately east of the avenue.[7] Density also rises with distance into Analco rather than concentrating at the boundary. I therefore do not interpret the density discontinuity as evidence of manipulation, although it leaves open the usual concern that residential sorting or boundary amenities may vary with land use.

Three features of the setting address other important sources of discontinuity, although they do not rule out residential sorting. First, the boundary precedes the houses. In 1542 the site was unurbanized, the settlements were placed on either side of the creek in one administrative act, and the city grew outward from them. Terrain, altitude, slope, climate, and hydrology are balanced across banks, so the creek did not separate different land (Section 2).

Second, the boundary is not a compound treatment (Keele and Titiunik, 2015). It coincides with no sub-municipal jurisdiction or school catchment,[8] the two banks share one municipality, court system, fiscal authority, police force, and labor market, and the Calzada connects them rather than dividing them (Figure A.9).

Third, present-day ethnicity does not provide an alternative partition at the boundary. In the 2020 census, 3.6% of Guadalajara's residents identify as Indigenous, 0.6% live in an Indigenous household, and 0.3% speak an Indigenous language, against national rates of 9.4 and 6.1% for the latter two. These low urban shares are consistent with the greater concentration of Mexico's

[7] The commercial core dates to the original spatial organization of the city described in Section 2.

[8] The Secretaría de Educación Jalisco organizes schools into administrative "*zonas escolares*," but these are used for inspection, planning, and seat allocation rather than for enforcing strict residential catchment rules.

Indigenous population in rural areas (Woo-Mora and Torres-López, 2026). Interviewer-rated skin tone in a 2017 mobility survey is likewise dispersed across the metropolitan area rather than clustered on either bank (Appendix A.7). The colonial assignment did bundle status with ethnicity, and the comparison between Analco and Mexicaltzingo helps separate them: both were Indigenous *pueblos*, but only Analco carried the excluded status. If ethnicity drove the gap, it should narrow along their border.

What remains testable. Several claims about present-day conditions remain. The gap may reflect contemporary housing prices and socioeconomic sorting rather than anything the assignment set in motion. Schools may still differ in supply or quality. Crime may offer different outside options to young people on each side. Each of these claims is testable, and I take them in turn.

Interpretation. The boundary separates locations by which side of the colonial hierarchy they were assigned to in 1542. Under the continuity assumption, Equation (1)'s τ captures the long-run effect of assignment to the excluded side on human capital today, rather than the effect of any single channel. It captures what accumulated on each side over five centuries, including the initial status, the differential provision that followed it, and the social differences that developed alongside. It is measured after public provision converged.

In the terms of Manski (1993), τ aggregates the channels of neighborhood influence rather than separating them. Jurisdiction, geography, labor-market access, and current school access do not change discontinuously at the boundary, so they do not provide competing discontinuous treatments. The design does not, however, separate place effects from the composition of residents on each side. Historical residential sorting may itself be part of the adjustment to the original assignment, but current sorting on unobserved determinants of human capital would threaten the continuity assumption. Endogenous and contextual effects therefore remain bundled inside the estimate.[9]

The two outcomes are formed at different times. Completed schooling reflects decisions taken across many decades, including cohorts schooled in the mid-twentieth century. A gap in completed schooling is therefore consistent with differences in educational provision that existed for earlier cohorts but later disappeared. Current attendance, by contrast, is a decision being made today. A gap in attendance therefore cannot be attributed only to differences in provision faced by older cohorts.

[9]This is the distinction in Chyn and Katz (2021) between exposure effects accumulated over childhood and the contemporaneous effect of the current environment. The stock relies on current residence proxying for past residence; the flow mixes the present environment with the child's accumulated exposure. Neither separates the present environment from the historical one. A parent's own schooling was itself shaped by the provision asymmetry of Section 4.1, and the attendance specifications condition on it, which holds the inherited stock fixed without removing whatever of that asymmetry reaches the child through the parent by other routes.

3.3 A boundary gap that persists across a century

3.3.1 Human capital gaps at the boundary in 2020

Panel (a) of Figure 2 shows a clear discontinuity in years of schooling at the historical boundary: residents immediately east of Calzada Independencia attain approximately 0.8 fewer years of schooling than residents immediately west, a gap equal to about 9% of the municipal mean, or 0.17 standard deviations; adjusting for predetermined covariates leaves a gap of 0.14 standard deviations. The discontinuity is visible within 100-meter bins on either side of the boundary.

Panel (b) turns to current school attendance among individuals aged 6–22, an age range chosen to ensure sufficient density of observations near the boundary. School attendance drops by about 6 percentage points at the boundary. Controlling for parental education, household size, and migration status absorbs the part of the gap that reflects the inherited human-capital stock and leaves a drop of approximately 3.5 percentage points, or 0.09 standard deviations, within meters of crossing into the Analco side (Table B.1). The gap therefore appears not only in completed schooling but also in current enrollment, among comparable households facing no discontinuity in school access or quality at the boundary.

Table B.1 reports the corresponding RDD coefficients in raw and standardized form across alternative polynomial orders, kernel weighting schemes, variance estimators, and specifications with and without additional individual covariates. The estimates are stable in sign and magnitude, with the standardized schooling effect lying between -0.08 and -0.17 standard deviations across every specification. Precision tracks the effective sample: the covariate-adjusted attendance specification retains about half the observations of the preferred one, and its confidence interval widens accordingly.

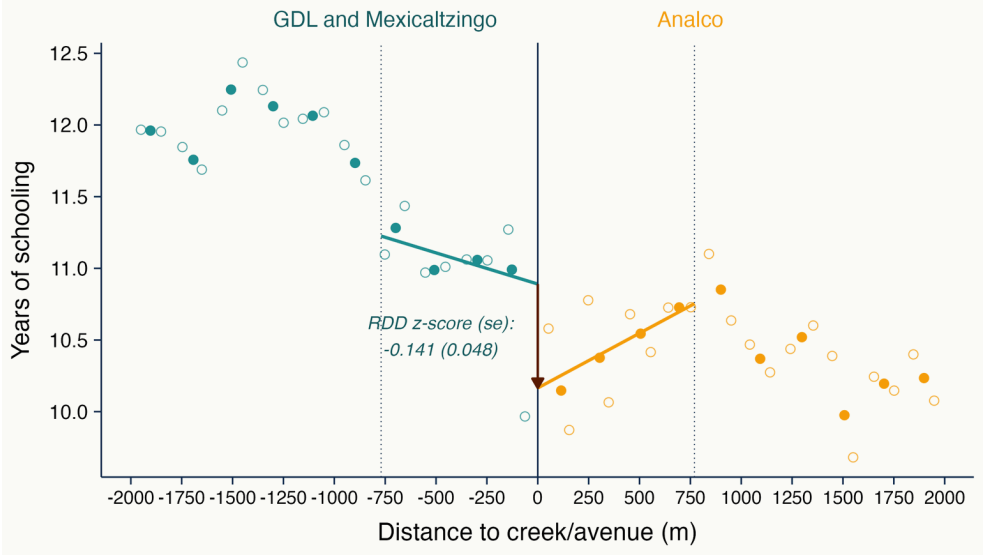
Housing-market sorting and capitalization. If housing prices capitalize local amenities or the historical stigma of the boundary, socioeconomic groups may sort across space, and the resulting gaps could reflect selection, as Baldomero-Quintana et al. (2025) document for Mexico City.

Using parcel-level cadastral data from the Guadalajara Municipal Cadaster, I find a clear west–east gradient in land values (Figure B.2a), which correlate strongly with online housing listings (Figure B.2b). Because cadastral values are recorded at the parcel level and aggregated to blocks, I re-estimate the discontinuity at the block level and control linearly for average land values.

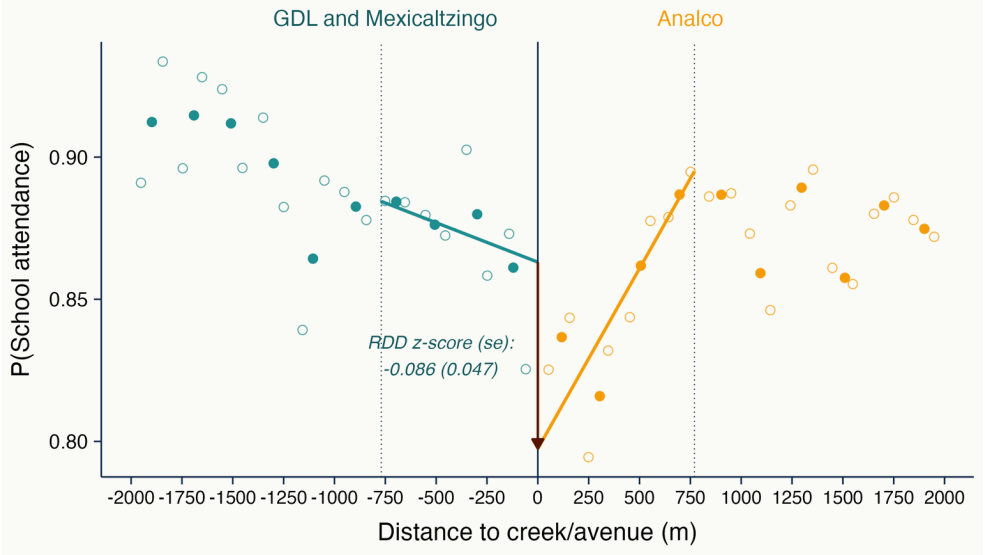
In the preferred specification, which includes clean controls and land values, the gap is 0.39 block-level standard deviations, equivalent to about 0.78 years of schooling given a standard deviation of 2.00 years across blocks (Table B.2). Across specifications controlling for land values, the gap ranges from 0.78 to 0.97 years, bracketing the 0.79-year estimate from the individual-level data.[10] Because land values may themselves capitalize the boundary's history, controlling for them may

[10] This check covers years of schooling only: the public block-level census suppresses blocks with few households, preventing the construction of a comparable measure of school attendance.

Figure 2: Spatial RDD estimates of human capital gaps at the historical boundary (2020)



(a) Years of schooling, ages 6+



(b) School attendance, ages 6–22

Notes: Binned means and local linear fits from individual-level data in the complete 2020 *Censo de Población y Vivienda* for the municipality of Guadalajara. The horizontal axis is the signed orthogonal distance in meters to the historical boundary, with zero denoting Calzada Independencia (the former San Juan de Dios creek); positive values are on the eastern (Analco) side. The vertical axes display the outcome in its original units: years of schooling (Panel a, individuals aged 6 or older) and the probability of school attendance (Panel b, individuals aged 6–18). Hollow circles represent mean outcomes within 100-meter bins; solid circles within 200-meter bins; binning is restricted to the window $[-2000, +2000]$ meters around the boundary. Vertical dotted lines mark ±770 meters, the bias-correction bandwidth b selected for years of schooling following Calonico et al. (2014); the thick colored lines are local linear fits estimated within that window without covariates, and the same window is used in both panels. The coefficients annotated in the figure are the covariate-adjusted standardized estimates reported in Table B.1, each computed at its own MSE-optimal estimation bandwidth h (395 meters in Panel a, 895 meters in Panel b), with sex, age, longitude, latitude, and boundary-segment fixed effects, and in Panel (b) parental education, household size, and migration status. The plotted fits and the annotated estimates therefore come from different specifications. No curve or band outside those fits is shown. A downward jump at the boundary indicates a relative human-capital disadvantage on the Analco side.

absorb part of the effect of interest rather than simply remove a confound.

School supply and quality. If the gap were sustained by unequal schooling, the boundary should still divide the schools themselves. Spatial RDD estimates find no discontinuity in the number of schools per block, nor in any of eight supply measures broken down by sector, level, and shift; every FDR-adjusted q-value equals unity (Table B.3).

Quality is a partial exception. Across *Formato 911* administrative records, schools on the eastern bank begin the series at a disadvantage. Through about 2002 they show lower approval rates, higher repetition and larger classes, with confidence bands away from zero. Those gaps close over the following few years, and from the mid-2000s onward none of the six indicators shows a systematic differential, in boundary RDDs or in municipality-wide comparisons (Figures B.4 and B.5). ENLACE standardized scores in Spanish and mathematics, which begin in 2006, show no differential across their whole span (Figures B.6 and B.7). The attendance gap in Table B.1 is measured in 2020 among people aged 6 to 22, who attended school after that convergence, so the early difference does not account for it. For completed schooling, which pools older cohorts, that difference is one of the channels the estimate aggregates rather than a confound to be netted out. Schools have no strict residential catchments, so a household on either bank can enroll a child across the avenue, a few hundred meters away. Measured supply at the line therefore understates the choice set available on both sides.

Crime exposure. Using over 190,000 georeferenced crime reports from the Jalisco Attorney General's Office (2017–2024), spatial RDD estimates show no significant discontinuities in overall crime or in any category at the boundary, including homicide, which is least subject to differential reporting (Table B.4 and Figure B.8). The boundary therefore does not coincide with a discontinuity in crime exposure.

3.3.2 The gap is largest where both sides were Indigenous

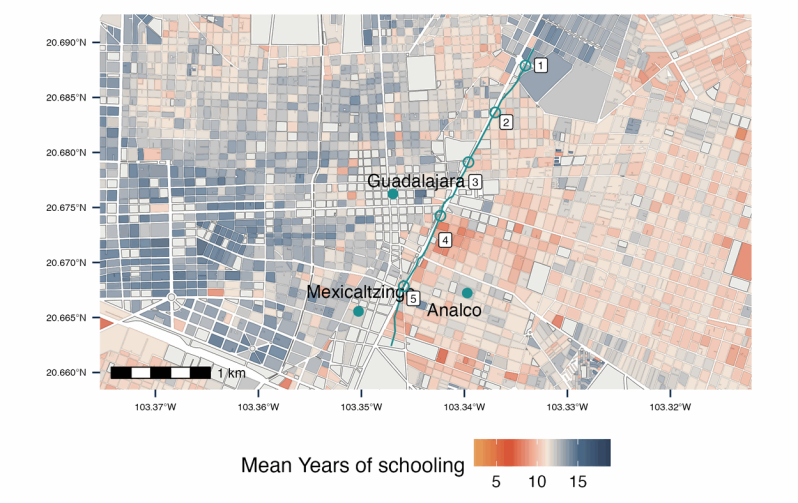
The colonial assignment bundled excluded status with ethnicity, and the boundary itself separates them. Along its northern stretch Analco faces the Spanish city; along its southern stretch it faces Mexicaltzingo, an Indigenous *pueblo* whose residents arrived as allies of the conquerors. If ethnicity drove the gap, it should narrow on the southern stretch, where both sides were Indigenous.

I estimate the discontinuity at each point along the boundary using the boundary-adaptive RDD of Cattaneo et al. (2025). The discontinuity is larger at each of the two southern points than at each of the three northern ones, for years of schooling and for school attendance alike (Table B.5).[11]

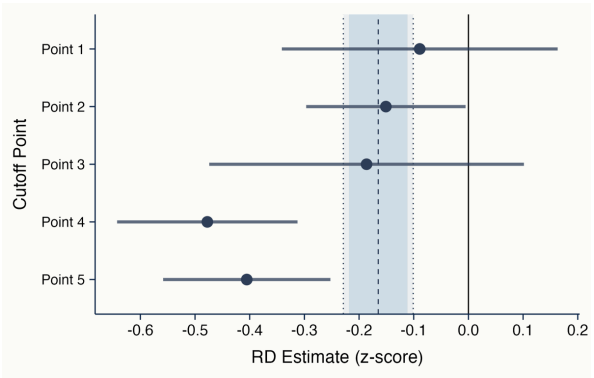
[11] Each of the two southern points is compared with each of the three northern ones, using the bias-corrected estimates and their standard errors. Five of the six contrasts are significant at the 5 percent level for each outcome and the sixth at the 10 percent level. Sharpened false-discovery-rate q-values (Anderson, 2008) computed across

The two southern points do not differ from each other, and the northern points do not differ among themselves. The gap is therefore widest where both banks of the creek held Indigenous *pueblos* and only one of them carried the excluded status.

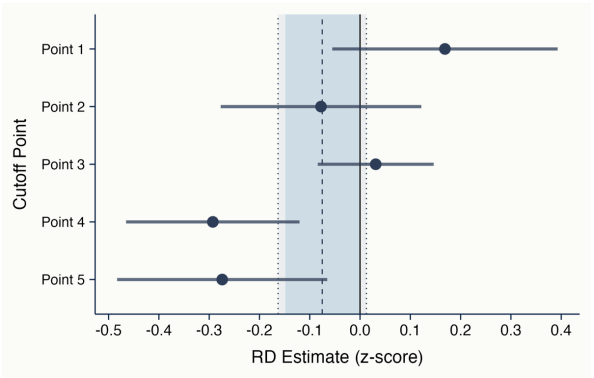
Figure 3: Boundary-adaptive RDD estimates (2020)



(a) Study zone near the historical boundary



(b) Years of schooling



(c) School attendance

Notes: Panel (a) shows the study zone around Calzada Independencia; Figure 1b shows the full municipality on the same color scale. Panels (b) and (c) report boundary-adaptive RDD estimates and uniform confidence bands following Cattaneo et al. (2025), using individual-level data from the 2020 *Censo de Población y Vivienda*. The aggregated estimate is shown at the bottom of each panel. The vertical dashed line marks the univariate RDD estimate from Figure 2; shaded areas show the 95% and 90% confidence intervals.

---

the six comparisons within each outcome stay below 0.03 throughout, so every contrast survives the correction for multiplicity. Because `rd2d` does not report the covariance between boundary points, and the version available when these estimates were produced offered no joint test of equality across points, the standard error of the difference treats them as independent. That overstates the standard error when the covariance between the two estimates is positive, which overlapping local neighbourhoods would ordinarily produce, but bias-corrected local-polynomial estimates carry influence weights that can change sign, so the direction is not guaranteed. These are pointwise comparisons rather than a simultaneous test over the six of them, and a sixth boundary point returns no bias-corrected estimate and is not covered.

3.3.3 The gap is not explained by income

I now estimate the same east–west gap parametrically using the 2015 Intercensal Survey.[12] The coefficient on the Analco indicator is the fitted gap at the boundary, projected from side-specific linear trends estimated on households across both banks. The survey also records household income, allowing me to ask whether present-day income accounts for the schooling differences.

These comparisons come at a cost. The estimate is exposed to the sorting and provision concerns that the local design addresses, and it depends on the assumed linear gradient. It draws on households across the municipality rather than only the blocks adjoining the creek. Table 1 reports the results.

Table 1: Human capital and income gaps in Guadalajara (2015)

	Years of schooling	$\mathbb{1}\{Attends\ school_i\}$							Monthly income (log)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\mathbb{1}\{Analco_i\}$	-1.31***	-0.074***	-0.052***	-0.081***	-0.054**	-0.051**	-0.052**	-0.015*	-0.084***	-0.010	0.075
	(0.118)	(0.017)	(0.015)	(0.024)	(0.024)	(0.023)	(0.025)	(0.008)	(0.027)	(0.024)	(0.068)
Years of schooling (YoS)										0.065***	0.068***
										(0.001)	(0.005)
$\mathbb{1}\{Analco_i\}\times$ YoS											-0.006
											(0.006)
Parental YoS					0.021***	0.018***	0.019***	0.003***			
					(0.001)	(0.001)	(0.001)	(0.0006)			
Household income (log)							0.0002	0.0000			
							(0.008)	(0.002)			
Sample (age in years)	6+	6–25	6–18	12–18	12–18	12–18	12–18	6–11	18+	18+	18+
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Household size						✓	✓	✓			
Observations	74,779	27,971	17,463	9,922	9,916	9,916	9,041	8,059	31,584	31,478	31,478
Dep. var. mean	9.18	0.685	0.884	0.807	0.807	0.807	0.803	0.984	8.65	8.65	8.65
Dep. var. SD	4.71	0.465	0.320	0.395	0.395	0.395	0.398	0.127	0.644	0.645	0.645
R^2	0.443	0.387	0.181	0.127	0.172	0.183	0.193	0.020	0.123	0.268	0.269

Notes: OLS estimates from individual-level data in the 2015 *Encuesta Intercensal* for the municipality of Guadalajara. The treatment indicator $\mathbb{1}\{Analco_i\}$ equals one if individual i resides east of the historical SJdD creek (today Calzada Independencia). *Individual controls* include sex and age fixed effects. *Geographic controls* include block longitude, latitude, log distance to the central business district, signed distance to the historical boundary, and the interaction of distance to the boundary with the Analco indicator. The first column reports years of schooling for individuals aged 6+. The next seven columns report the probability of attending school for the indicated age range. The fifth through eighth columns additionally control for the maximum of mother's and father's years of schooling; the sixth through eighth for household size; and the seventh and eighth for log household income. The final three columns use log monthly income (2015 MXN) as the dependent variable and restrict the sample to adults aged 18+. Standard errors clustered at the household level are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

The first column shows that residents east of the boundary have, on average, 1.31 fewer years of schooling than residents west of it. Against a municipal mean of 9.18 years, which sits just above the nine years of compulsory basic education, this is a gap of about 14% of the mean, or close to half of the three-year lower-secondary cycle.

The next three columns show that the gap is also present in current school attendance. Attendance

[12] The Intercensal sample thins too quickly near the boundary to fit a local polynomial within a data-driven bandwidth. Identification rests instead on an assumed linear gradient in distance over a window spanning five kilometers on either side of the boundary, which covers most of the municipality.

is about 11% lower on the east side among individuals aged 6–25, 6% lower among ages 6–18, and 10% lower among adolescents aged 12–18.

The next three columns add controls for parental education, household size, and household income. Parental schooling absorbs about a third of the adolescent gap, which falls from 8.1 to 5.4 percentage points; household size and current income change it little thereafter, leaving 5.2 percentage points, or about 6.5% of the mean attendance rate. Parental schooling is itself an outcome of the historical assignment, so the attenuation indicates how much of the adolescent gap runs through the previous generation's education rather than identifying a confound. Current household income does not account for the gap that remains. Among children aged 6–11, the gap is smaller at 1.5 percentage points, consistent with near-universal primary attendance.

The final three columns turn to income. Adults east of the historical boundary earn about 8.1% less per month than adults west of it. Once years of schooling enter as a control, the income gap falls to 1% and is statistically insignificant. The interaction between the Analco indicator and years of schooling is also close to zero and statistically insignificant. Consistent with Guadalajara forming a single local labor market (Aldeco et al., 2024), returns to schooling are similar across the two banks. The income difference therefore suggests differences in human capital rather than different returns to schooling.

The historical boundary therefore predicts the spatial distribution of schooling across the municipality, as Figure 1b shows. The same east–west pattern continues beyond Guadalajara's municipal border and into neighboring municipalities of the metropolitan area (Figure B.9). The persistence of this spatial pattern is consistent with path dependence in the city's development, with the colonial assignment continuing to shape the urban equilibrium long after the original institution disappeared (Lin and Rauch, 2022; Hanlon and Heblich, 2022).

3.3.4 The gap is already present in 1930

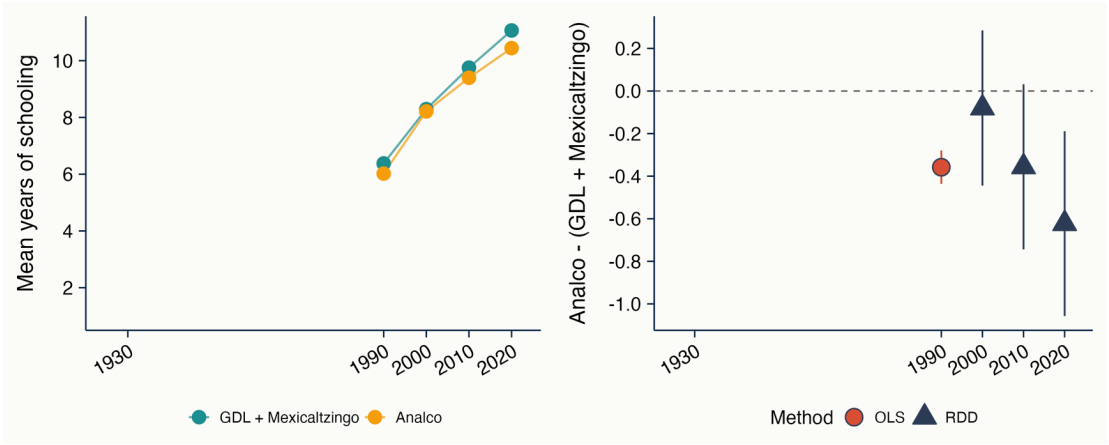
The earliest geolocated censuses make it possible to ask how far back the boundary gap can be traced. I use the 1930 census at the *cuartel* level, the 1990 census at the AGEB level, and the 2000, 2010, and 2020 censuses at the block level.[13] Spatial resolution improves across waves, with the 2000–2020 censuses allowing the discontinuity to be estimated at the same fine scale as the main results.

The gap is already visible in the earliest wave. In the 1930 census, the first with individual-level records on literacy and school attendance for Guadalajara (Anderson and Spike, 2007), adult illiteracy was 43.2% on the Analco side and 31.4% on the west among *cuarteles* adjacent to the boundary.[14] Among individuals aged 18 or younger, a group that includes children below school-

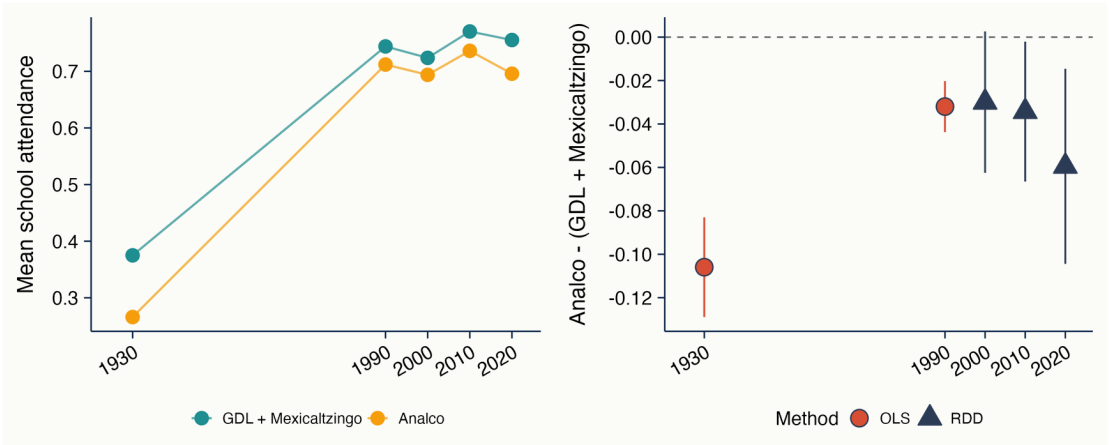
[13] For 1930 and 1990, I compare *cuarteles*/AGEBs immediately adjacent to the boundary using a parametric specification. For 2000–2020, I estimate the non-parametric spatial RDD of Equation (1) at the block level, matching observations to the boundary using the same georeferencing procedure described above.

[14] Analco comprises *cuarteles* 8 and 9 on the east; central Guadalajara and the historical Mexicaltzingo area

Figure 4: Human capital gaps across the historical boundary over time (1930–2020)



(a) Years of schooling



(b) School attendance

Notes: Evolution of the boundary gap in (a) years of schooling and (b) school attendance. Each panel shows, at left, the mean outcome on each side of the boundary, with the western *cuarteles* covering Spaniard Guadalajara and Mexicaltzingo and the eastern side covering Analco, and, at right, the Analco-minus-west gap by census wave. Circles are OLS estimates for 1930 (*cuartel* level) and 1990 (AGEB level); triangles are non-parametric spatial RDD estimates for 2000–2020 at the block level following Calonico et al. (2014). Vertical bars show 95% confidence intervals. All estimates use individual-level census microdata for the municipality of Guadalajara. Panel (a) has no 1930 estimate because the census does not record years of schooling in that year; 1930 is retained on the axis for comparability with panel (b). The 2000–2020 estimates are uncontrolled non-parametric RDD contrasts using a uniform kernel and the Calonico et al. (2014) optimal bandwidth. The 1930 and 1990 estimates are parametric contrasts with geographic controls. For years of schooling, the 1990 estimate is measured among adults aged 18+, whereas the 2000–2020 estimates cover ages 6+, so the pre-2000 estimates are not directly comparable with the block-level series. The uncontrolled RDD contrasts also differ from the covariate-adjusted estimates in Section 3.3.1: in 2020, the schooling gap is about −0.62 years without covariates and about −0.8 years with covariates, while the corresponding attendance gaps are about 6 and 3.5 percentage points. Adult literacy is reported in Figure B.10.

entry age, school attendance was 26.6% in Analco and 37.5% on the west. Controlling flexibly for age and sex leaves a gap of 10.6 percentage points, against the 10.9-point raw difference, so the contrast does not rest on differences in age composition across banks (Table B.6). The educational divide was therefore already present nearly a century ago, before the major local and federal expansions of public schooling in Mexico.

Figure 4 traces the divide in later censuses. Because spatial units and samples differ before 2000, the 1930 and 1990 estimates establish historical persistence, while the comparable 2000–2020 block-level estimates trace the more recent evolution. By 1990, adults on the Analco side had about 0.36 fewer years of schooling. The block-level schooling gap is close to zero in 2000, about 0.36 years in 2010, and about 0.62 years in 2020.[15] The attendance gap is about 3 percentage points in 2000 and 2010 and about 6 percentage points in 2020. Adult literacy divided the two banks sharply in 1930, at 68.6 percent on the west against 56.8 on the east, and that margin closed as literacy approached universal coverage; it shows no discontinuity at the boundary in any wave from 1990 onward (Figure B.10).

The boundary therefore marked an educational divide by 1930, and that divide remains visible today despite the subsequent expansion of public education and the convergence in measured school supply and quality across the two sides.

3.3.5 Robustness checks

Beyond the convergence results, I subject the spatial RDD estimates using the 2020 census to several diagnostic checks: alternative estimators and spatial aggregations, residential turnover, spatial autocorrelation, and the possibility of spurious discontinuities. The estimated schooling gap remains consistent in magnitude. Full results are in Appendix B.

Migration and residential stability. According to the 2020 census, 91.5% of Guadalajara residents lived in Jalisco five years earlier; re-estimating the RDD excluding inter-state migrants yields nearly identical results (Table B.1). The gap is therefore unlikely to reflect an inflow of recent, less-educated arrivals into Analco. A proxy for residential persistence, built by matching 2020 adults to demographically similar 2010 residents of the same block, points the same way, but the census is anonymous and the match distance falls mechanically with the number of records available, so Appendix B.12 does not read it as a measured difference in turnover. Turnover volume is in any case not selection, and the census says nothing about the schooling of former residents, so none of this bounds the selection of those who left.

comprise *cuarteles* 1, 2, and 7 on the west.

[15] The near-zero estimate in 2000 may partly reflect the 1992 Reforma gas explosions, which destroyed and repopulated part of Analco within the RDD bandwidth shortly before the census. Figure 10 shows that the rebuilt blocks were more educated than the rest of Analco in 2000 and converged afterward. Part of the increase between 2000 and 2020 may therefore reflect the fading of this temporary disruption.

Donut estimates. Residential density at the boundary is far lower on the western bank, so the innermost observations come disproportionately from one side. Dropping every resident within 50, 100 or 150 meters of the line, estimating without covariates throughout, leaves the schooling gap between −0.10 and −0.21 standard deviations and the attendance gap between −0.09 and −0.13 (Table B.7).

Spatial autocorrelation and placebo boundaries. Following Müller and Watson (2024), and using Becker et al. (2025) package, I reject a spatial unit root and fail to reject spatial stationarity, supporting local RDD estimators without spatial differencing. Finally, randomly shifting and rotating the boundary 1,000 times within the municipality (Figure B.11) yields a placebo distribution tightly centered at zero, with the true-boundary estimate in the left tail; following Borusyak and Hull (2023), a recentered specification makes the true-boundary discontinuity *larger*, not smaller (Figure B.12).

4 Exclusion and Persistent Neighborhood Differences

I examine the two dimensions introduced in Section 2. I document historical differences in formal public institutions using the entries named on the city's general maps between 1800 and 1908, and relational differences using naming patterns in nineteenth-century records and present-day survey data.

4.1 Historical exclusion in formal public institutions

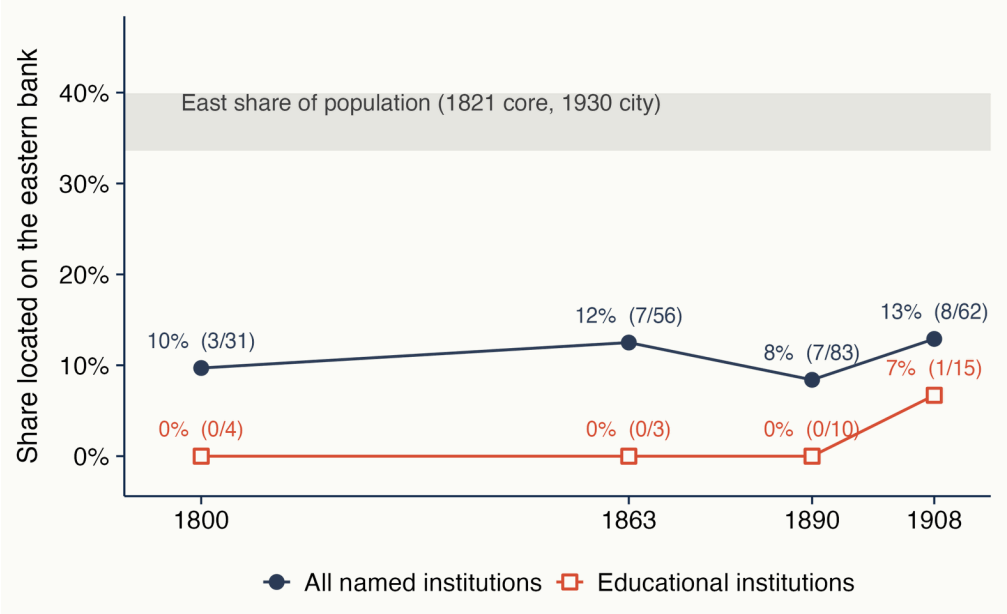
Four general maps of Guadalajara name the city's institutions and locate them on the urban grid: 1800, 1863, 1890, and 1908. I georeference each map and assign each located entry to a bank of the creek.[16] Figure 5 reports the share of entries located on the eastern bank in each wave, against the eastern share of the population. The eastern bank accounted for between a third and two-fifths of the population in the available enumerations, but for less than fifteen percent of the named institutions in every map.

No educational institution appears east of the creek in 1800, 1863, or 1890, and the first to appear, in 1908, is a girls' school run by a religious order. These indices record the institutions their authors considered worth naming, so they establish where the city's named educational establishments stood rather than the number of school places available on each bank.

Newly digitized parcel-level microdata from Guadalajara's 1882 cadaster show the same asymmetry in a market outcome: assessed land values are roughly half to two thirds lower on the eastern bank at the creek, decades before twentieth-century public schooling and industrialization (Appendix A.8).

[16] Appendix A.9 documents the sources, procedure, and counts by category.

Figure 5: Institutions located on each bank, 1800–1908



Notes: Each point is the share of the entries in a map's index that are located east of the SJdD creek, with the count of eastern entries over the total located in parentheses. The upper series covers all named institutions; the lower series covers educational establishments only. Each map reflects its own conventions about what merits an entry, so shares are comparable across banks within a wave, and across waves within a category present in all four, as the educational category is. The shaded band spans the eastern share of the population in the two enumerations that record residence by *cuartel*: 39.9 percent in the 1821 *padrón*, restricted to the central core for which sides can be assigned, and 33.6 percent in the 1930 census. Sides for the map entries follow the same boundary used in the spatial RDD; sides for the *cuarteles* follow the digitized blocks of the 1890 plan, in which *cuarteles* 8 and 9 lie east of the creek and 1–7 lie entirely west. Of the 105 entries in the 1908 index, 62 could be assigned to a bank, 24 by direct location on the plan and 38 by name correspondence with an earlier georeferenced wave; the remaining 43 are excluded. Appendix A.9 reports the sources of the four maps, the location method for every entry, and the counts by category.

Contemporary school provision shows no such asymmetry. As reported in Section 3, spatial RDD estimates find no discontinuity at the boundary in the number of schools per block or in any of eight supply measures (Table B.3), and administrative records covering 1998–2015 show no differential in school quality, either at the boundary or in municipality-wide comparisons. The supply comparison is local to the boundary, while the historical indices locate institutions across the two banks as a whole, so the two are measured on different spatial scales. Historical records and present-day surveys, by contrast, document relational differences across the same boundary.

4.2 Social differences across the boundary in historical records

Following recent work (Fryer and Levitt, 2004; Abramitzky et al., 2020; Bazzi et al., 2020; Weigand et al., 2025; Kamel, 2026), I use names as a proxy for local identity and cultural change: they are chosen by parents, transmitted within families, and, in this setting, tied to local religious devotions.

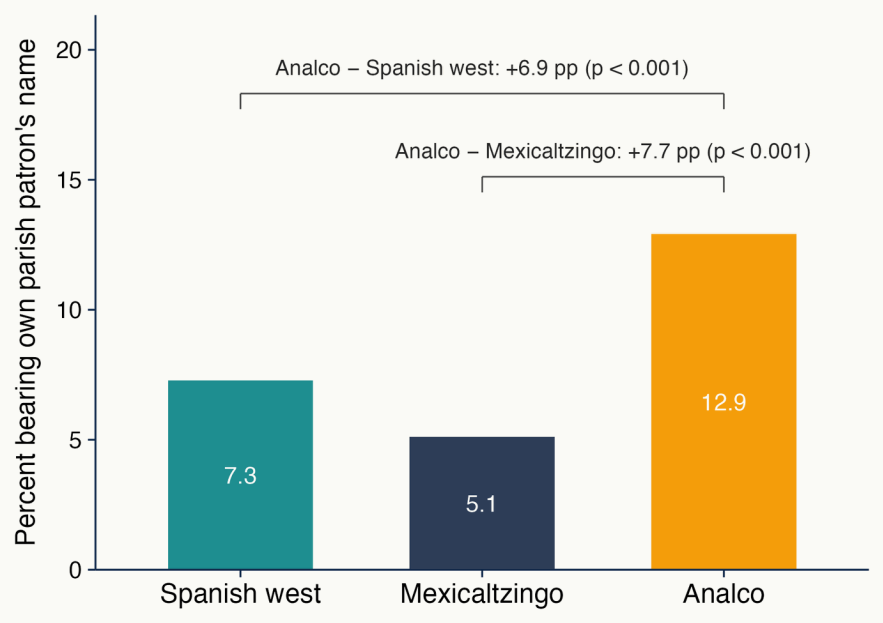
I use two historical sources. First, I use the civil *padrones* of 1821–22 and 1824 compiled by Anderson and Spike (2007), which record residents' given names, ages, and *cuarteles*. I complement them with the national 10 percent sample of the 1930 census, which records the same fields. Second, I use FamilySearch parish registers of baptisms and marriages spanning 1850–1949.

Naming after the local patron saint. Naming after the neighborhood's own patron saint was more common in Analco than in either of the two non-excluded settlements (Figure 6). I compare the share of men whose given name matches the patron saint of their own parish, scoring each *cuartel* on its own patron and comparing the same name token across settlements. Name-token, twenty-year cohort, and surname fixed effects absorb differences in the overall popularity of common saints' names, so the comparison isolates whether a token matches the local patron rather than how common the name is.[17]

In 1821, 12.9 percent of Analco men bore their parish patron's name, against 7.3 percent in the Spanish west, an adjusted difference of 6.9 percentage points. Analco also exceeds Mexicaltzingo, where 5.1 percent of men bore the parish patron's name, an adjusted difference of 7.7 percentage points. Mexicaltzingo itself is not distinguishable from the Spanish west. Analco and Mexicaltzingo were both Indigenous settlements, but only Analco carried the excluded colonial standing, so a difference in Indigenous status alone is difficult to reconcile with the higher patron naming in Analco.

[17]The eastern bank contains two parishes with distinct patrons, Analco under San José and San Sebastián and San Juan de Dios under San Juan. Pooling both eastern parishes against the western *cuarteles* with a known patron yields a 3.9 percentage-point east–west difference on 9,795 token–person observations. Restricting the comparison to a single eastern patron would make every José recorded in Analco match by construction; the 1824 enumeration covers neither San Juan de Dios nor Mexicaltzingo and therefore does not admit these settlement comparisons.

Figure 6: Naming after the local patron saint, by settlement, 1821



Notes: The figure plots the share of men whose given name matches the patron saint of their own parish in the census of 1821 (Anderson and Spike, 2007), for three settlements: the Spanish west (the western *cuarteles* with a known parish patron), Mexicaltzingo (*cuartel* 10, under San Juan), and Analco (*cuartel* 8, under San José and San Sebastián). The parish map is Figure C.1 in Appendix C. Bars report unadjusted own-patron shares in percent. The brackets report the difference between Analco and each comparison settlement from a linear probability model at the token–individual level, weighted by the inverse of the number of given-name tokens, with name-token, twenty-year cohort, and surname fixed effects; standard errors are clustered by household. Analco exceeds the Spanish west by 6.9 percentage points (95 percent CI [5.8, 8.1]) and Mexicaltzingo by 7.7 percentage points (95 percent CI [5.2, 10.2]); Mexicaltzingo and the Spanish west do not differ (−0.7 percentage points, 95 percent CI [−2.6, 1.3]). Figure C.2 shows the calendar-saint placebo: naming after the saint of the baptism day is equally common on both sides.

Two checks show that the result is specific to the neighborhood patron. First, it is not driven by religious naming in general. In the baptismal registers, children on both banks are about four percentage points more likely to be named after the saint of the baptism day than under a shifted-calendar placebo, with no difference across the two sides (Figure C.2).

Second, it is not driven by stronger family-name transmission on the east. Sons on the western side bear their father's given name at roughly twice the eastern rate in the 1821–22 and 1824 censuses and in the baptismal registers of 1850–99.[18]

The east–west difference is therefore specific to naming tied to the local parish, rather than to religious naming or family-name transmission more generally.

Distinct names and surnames. The difference extends beyond patron-saint names. For each census, I compute a distinctiveness index in the spirit of Fryer and Levitt (2004): the probability that a resident lives east of the creek given only their first name, estimated leave-one-out and

[18]The gap holds in the marriage acts of 1850–1949. Acts before 1850 are too sparse to support the comparison.

benchmarked against a permutation null.[19]

The first panel in Figure 7 shows a stable east–west gap for cohorts born between 1760 and 1839, with the 1821–22 and 1824 censuses agreeing cohort by cohort. The final bin of each enumeration is right-truncated by its census date, so the 1820–39 cohort is observed only through its earliest years in the two early censuses and the agreement across enumerations is exact only for the bins observed in full. In the 1930 census, the gap remains visible for the cohort born in 1860–79 but disappears for cohorts born after 1880, coinciding with the period in which Guadalajara's industrial expansion reorganized its social geography.

The first-name gap is not explained by ethnic composition. The 1821 census records ethnicity (*calidad*), and estimating the comparison within *calidad* cells yields the same pattern.

Given names converged, but surnames remained distinct. The second panel of Figure 7 repeats the index on paternal surnames, which capture residential lineage rather than parental naming choice. It is larger than the first-name index in the early censuses and, in 1930, rises across cohorts as first-name distinctiveness disappears, peaking in the cohort born in 1920–39. By 1930, first names no longer distinguished the two sides, but paternal surnames still did. The naming custom itself still differed in 1900–49: in the marriage acts, 7.7 percent of western grooms bore their father's name against 5.9 percent of eastern ones. The given-name repertoire converged while the two populations remained residentially separate.

4.3 Social differences across the boundary today

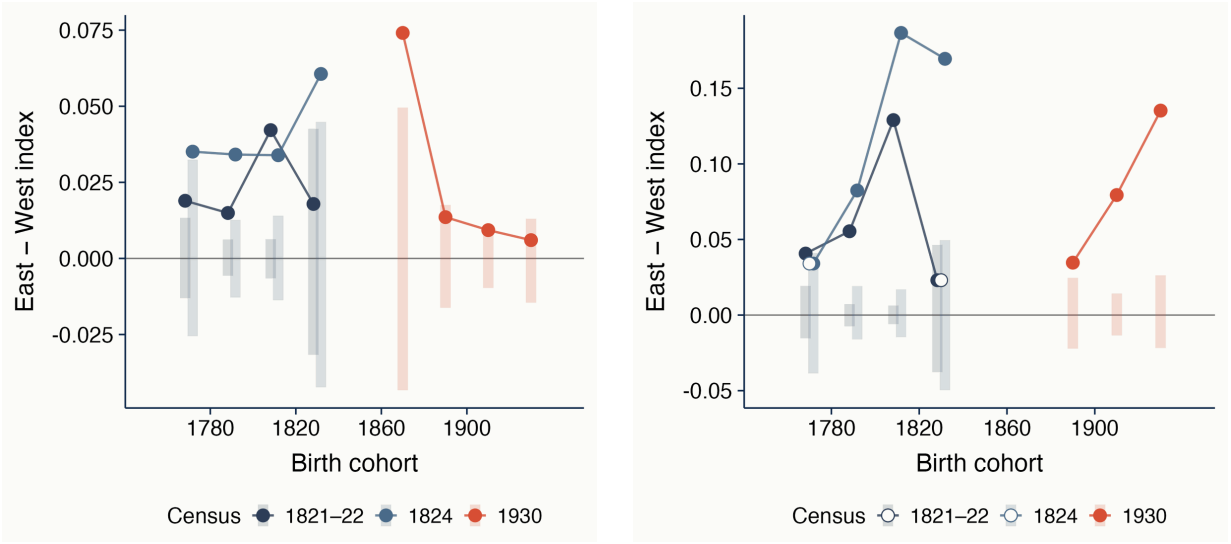
I now turn to the present and look at who people within Guadalajara trust, one of the several components of *social capital*. Durante et al. (2025) separate four of them: social participation, political participation, interpersonal trust, and trust in institutions, and find them only weakly correlated with one another.

I estimate the same parametric specification used in Section 3.3.3 on individual-level data from the 2022 and 2024 waves of a local Life Quality Survey, geolocated at the census-tract (AGEB) level.[20] The Life Quality Survey measures three of the four social capital dimensions: interpersonal trust, institutional trust, and associational participation. Given the number of outcomes, I report sharpened false-discovery-rate q-values (Anderson, 2008).

[19] Spellings are normalized before any comparison (Josef and José are treated as the same name written by different enumerators), and the index is estimated within census, so it never compares naming conventions across enumerators from different censuses.

[20] The Life Quality Survey is designed and produced by the local NGO *Jalisco Cómo Vamos* (JCV). All specifications include age-by-sex and survey-year fixed effects and control for the respondent's signed distance to the boundary interacted with the Analco indicator, so comparisons allow side-specific distance gradients; regressions on trust and membership additionally absorb years of schooling, so the estimated differences are not mechanically driven by the education gap.

Figure 7: Given-name and surname distinctiveness by birth cohort (1740–1939)



(a) Given names

(b) Paternal surnames

Notes: Panel (a) plots, by twenty-year birth cohort, the east–west difference in the mean leave-one-out probability that a resident lives east of the creek given their first name, estimated separately within the censuses of 1821–22, 1824, and 1930. Shaded bars are 95 percent permutation bands under random assignment of side within cohort. The earliest cohort of each census is omitted (born 1740–59 in 1821–22, and 1840–59 or 1860–79 in 1930, depending on the panel): the cells are small and the permutation bands wide. Cohorts born in 1760–1819 are observed in full in two independent enumerations; the 1820–39 bin is right-truncated in each census by its enumeration date, so its year and age composition differs across censuses. Sides follow the *cuartel* recorded for each household, verified against the manuscripts' street lists; in 1930, *cuarteles* 8, 9, and 11 are assigned to the east and 1–7 to the west, where *cuarteles* 8 and 9 lie east of the creek in the digitized blocks of the 1890 plan and *cuartel* 11, which that plan does not cover, holds under one percent of the sample. Spellings are normalized within census before comparison. Panel (b) repeats the exercise using paternal surnames; hollow points are not significant at the five percent level. Levels are comparable within, not across, censuses.

Interpersonal trust. The first column in Table 2 documents the gap in years of schooling, consistent with that found in the census data.

The second through ninth columns characterize interpersonal trust conditional on schooling. On a 0–1 scale, residents on the excluded side score 25.2 percentage points higher on the composite interpersonal-trust index in the second column, driven primarily by a 38.7 percentage-point difference in trust in neighbors in the third. The fourth through sixth columns show that trust in family and in coworkers or classmates is also higher, whereas trust in people encountered on the street is unchanged.

People on the east side have a higher "neighbor premium", as shown in the seventh through ninth columns. Trust in neighbors relative to bridging contacts, family, and nonlocal contacts is higher and statistically significant on the east side. The pattern therefore does not reflect generalized trust. Instead, trust is disproportionately concentrated within the immediate neighborhood. The gap appears across age brackets, including respondents born after the early 1990s (Figure C.3).

Table 2: Localized social capital and educational attainment

	Years of schooling (1)	Interpersonal trust index (2)	Trust in neighbors (3)	Trust in family (4)	Trust in coworkers/ classmates (5)	Trust in people on the street (6)	Neighbor vs. bridging trust (7)	Neighbor vs. family trust (8)	Neighbor vs. nonlocals trust (9)	Years of schooling (10)
$\mathbb{1}\{Analco\}$	-1.95**	0.252***	0.387***	0.140*	0.217**	0.004	0.098**	0.123*	0.093**	-0.864
	(0.795)	(0.066)	(0.104)	(0.078)	(0.104)	(0.090)	(0.045)	(0.064)	(0.044)	(0.861)
Trust in neighbors										0.714
										(0.473)
Trust in neighbors × Analco										-1.68**
										(0.662)
Observations	578	578	578	578	578	578	578	578	578	578
Dep. var. mean	11.34	0.5944	0.6678	0.9135	0.5848	0.2180	0.6298	0.3772	0.6003	11.34
R^2	0.2124	0.1123	0.1234	0.1031	0.0838	0.1285	0.0727	0.1158	0.0961	0.2207
q-value		[0.001]	[0.001]	[0.057]	[0.035]	[0.288]	[0.035]	[0.045]	[0.035]	
Age-Sex fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Years of schooling fixed effects		✓	✓	✓	✓	✓	✓	✓	✓	

Notes: OLS estimates from the 2022 and 2024 waves of the local Life Quality Survey (*Jalisco Cómo Vamos*) for the municipality of Guadalajara. The treatment indicator $\mathbb{1}\{Analco_i\}$ equals one if the respondent's geolocated census tract lies on the eastern (Analco) side of the historical SJdD creek. Column (1) reports the raw Analco gap in years of schooling. Columns (2)–(9) report the Analco gap across interpersonal-trust dimensions, each net of years of schooling (absorbed by fixed effects). Column (10) regresses years of schooling on trust in neighbors, its interaction with the Analco indicator, and the Analco indicator: the negative neighbor-trust/schooling association is concentrated on the excluded side. All specifications include age-by-sex and survey-year fixed effects and control for distance to the boundary. Heteroskedasticity-robust standard errors in parentheses. Sharpened FDR q-values (Anderson, 2008), adjusted across all social-capital outcomes, in brackets.

Institutional trust. The eastern advantage in neighbor trust is accompanied by a pronounced deficit in confidence in formal institutions. Residents east of the boundary score 24.7 percentage points lower on the institutional-trust index (about the size of the sample mean), with lower confidence in every institution surveyed, including the judiciary, government, political parties, and the police. The first column in Table C.1 reports the index. High trust in proximate ties therefore coexists with low confidence in institutions outside the neighborhood. Part of the police gap may reflect exposure: the eastern side hosts the Calzada corridor and the San Juan de Dios market.

Participation. The same asymmetry appears in civic participation. Residents on the eastern side are 8.5 percentage points more likely to belong to a neighborhood or residents' association (about twice the sample mean) but are no more likely to participate in other types of organizations (Table C.2). Civic engagement is thus denser within the neighborhood but not broader beyond it.

The Life Quality Survey therefore documents higher neighbor trust and neighborhood participation on the eastern side, alongside lower institutional trust. These differences are descriptive.

4.4 Neighborhood orientation and schooling

The material dimension of exclusion is visible in the historical record and no longer at the boundary today. The relational dimension is visible in both.

Relational differences associated with the colonial divide were already visible by the end of the colonial period and remain visible in different forms today. The historical records and present-day surveys document these different expressions of local orientation on the eastern side. I refer to this pattern as *neighborhood orientation* (Dickinson, 1961; Carmon, 2001).

Today, trust and participation are more concentrated within the neighborhood on the eastern side. Historically, naming was more closely tied to the local parish, a central neighborhood institution, and given names and surnames were more distinctive across the same boundary. These measures characterize neighborhood orientation in different periods; they do not measure a single latent trait continuously over time.

This orientation is also related to schooling. The tenth column of Table 2 shows that, on the western side, neighbor trust is positively but imprecisely associated with schooling. On the historically excluded side, the interaction coefficient turns this relationship negative: higher trust in neighbors is associated with lower educational attainment.

The association is observational and does not establish whether neighborhood orientation affects educational investment or whether schooling changes neighborhood ties. Appendix C.4 sets aside two accounts of it, finding the gap unrelated to the schooling already attained in the surrounding tract and unrelated to land values across the boundary, so it is neither conformity to the attainment of neighbors nor wealth-based sorting. That does not identify a channel. Testing whether making the neighborhood salient affects educational investment requires exogenous variation in salience at the moment of choice.

5 Neighborhood Orientation and Educational Demand

The evidence so far documents relational differences across the historical boundary, but does not show whether those differences affect educational choices. I address that question with a pre-registered incentivized choice experiment that randomizes whether the neighborhood is salient when households choose an educational input. A simple framework clarifies what this variation can reveal.

5.1 A static framework

Let $e_i \in \{0, 1\}$ denote educational investment by agent i in neighborhood ℓ. Agents differ in whether local social considerations enter the decision. I represent this with $\theta_i \in \{0, 1\}$. Type-1 agents place weight on local social relations and on conformity with local educational behavior; type-0 agents do not. Let q_ℓ denote the share of type-1 residents and $\bar{e}_\ell$ average educational investment in the neighborhood.

Utility is

$$U_i(e_i \mid \theta_i, q_\ell, \bar{e}_\ell) = w_0 + \big[\rho - C(P_\ell)\big]e_i + a(s)\theta_i \left[\Gamma q_\ell(1 - \delta e_i) - \kappa|e_i - \bar{e}_\ell|\right]. \tag{2}$$

The term $\rho > 0$ is the private return to education. Public educational provision P_ℓ affects the effective cost of educational investment, $C(P_\ell)$, with $C'(P) < 0$. I interpret this cost broadly enough to include access, availability, travel, direct costs, and other frictions associated with weak educational infrastructure. Provision affects $C(P_\ell)$, not ρ.

The local component has two parts. The first, $\Gamma q_\ell(1-\delta e_i)$, captures the value of local interaction and the share δ of that payoff crowded out by educational investment. The second, $\kappa|e_i - \bar{e}_\ell|$, captures the cost of departing from local educational behavior. The distinction parallels the contextual and endogenous social-interaction effects in Manski (1993). The parameter $a(s)$ gives the weight placed on this local component when the neighborhood has salience s.

For a type-1 agent,

$$U_i(1) - U_i(0) = \rho - C(P_\ell) - a(s)\left[\delta\Gamma q_\ell + \kappa(1 - 2\bar{e}_\ell)\right]. \tag{3}$$

Define

$$\Pi_\ell \equiv \delta\Gamma q_\ell + \kappa(1 - 2\bar{e}_\ell) \tag{4}$$

as the *neighborhood pull*. A type-1 agent invests if and only if

$$\rho - C(P_\ell) \geq a(s)\Pi_\ell. \tag{5}$$

Provision and salience affect different margins. Better provision lowers the effective cost of education on the left-hand side. Salience changes the weight placed on local considerations on the right-hand side. When $\Pi_\ell > 0$, greater neighborhood salience reduces the relative payoff to educational investment. When the pull is small, the payoff shift is small and, with heterogeneity across households, few choices cross the investment margin; when $\Pi_\ell < 0$, greater salience can instead favor education.

The experiment randomizes neighborhood salience, not q_ℓ, $\bar{e}_\ell$, or any component of Π_ℓ. In the model, a change from $a(0)$ to $a(1) = a(0) + \Delta a$ changes the type-1 utility difference by

$$\Delta\big[U_i(1) - U_i(0)\big] = -\Delta a\, \Pi_\ell. \tag{6}$$

Empirically, I estimate the intention-to-treat effect of assignment to the salience treatment on educational demand. The experiment does not separately identify Δa, Π_ℓ, θ_i, q_ℓ, Γ, or κ.

The full model is in Appendix D. There, unequal historical provision affects education through $C(P_\ell)$, and the dynamic block shows how education and neighborhood orientation can subsequently reinforce one another. The state q_ℓ is a theoretical object: the historical naming records and present-day surveys document different expressions of local orientation, not repeated measurements of a single latent q_ℓ.

5.2 Incentivized choice experiment

5.2.1 Design

I test whether making the neighborhood salient changes educational demand in a pre-registered incentivized choice experiment around the historical boundary. Enumerators interviewed parents

and primary caregivers living within approximately two kilometers of the boundary and with at least one child aged 6–20.[21]

The analysis uses 239 households interviewed through July 2026, 193 east of the boundary and 46 west, close to the pre-registered target of 200 and 50. The unequal allocation was pre-specified to increase power on the historically excluded side.

Sampling. Sampling was stratified by side of the historical boundary. Within each side, I selected census blocks containing residents aged 6–20 in the 2020 Population Census, and enumerators screened every dwelling in the selected blocks. Every eligible household contacted participated.[22] The mean distance to the historical boundary is 1.5 km and the median is 1.7 km.

Neighborhood salience. The treatment changes the order of two tasks: a neighborhood-mapping exercise and the educational choice.

In the mapping task, respondents viewed a satellite map centered on their residence and traced the area they considered their neighborhood (*colonia* or *barrio*), following McCartan et al. (2024). They also reported what guided the boundary they drew.

In the *Map-First* arm, respondents mapped their neighborhood immediately before the educational choice. In the *Map-After* arm, they did so immediately afterward. Assignment was randomized within side of the historical boundary. The price-list instructions, the practice round, and the comprehension checks precede the randomization and are therefore identical in both arms; Figure E.1 shows the full order of tasks. The mapping exercise mentions neither education nor the historical boundary. The treatment therefore changes whether respondents have just thought about their neighborhood when they make the educational choice, while holding the information supplied by the survey fixed.

Educational demand. I elicit willingness to pay for one hour of online mathematics tutoring for one of the respondent's children. The elicitation uses a multiple-price list following Jack et al. (2022). Respondents make eleven choices between the tutoring session and an electronic Amazon gift card whose value ranges from MXN 0 to MXN 250 in MXN 25 increments. I randomize the order of the rows and whether tutoring appears on the left or right, and control for both versions in the regressions.

One of the eleven rows is randomly selected to bind, and one in ten households is selected for payment. A selected household receives either the tutoring session or the electronic gift card

[21] Fieldwork began on May 9, 2026 and was carried out by Descifra Investigación Estratégica using the SurveyToGo platform. The study was approved by the Paris School of Economics Institutional Review Board (approval 2026-015) and pre-registered in the AEA RCT Registry (AEARCTR-0018600).

[22] All participating households received MXN 100 regardless of their experimental choices. The participation payment was introduced on May 13, 2026; the 17 households interviewed before that date were compensated retroactively. The results are unchanged when these households are excluded.

implied by its choice in the binding row.[23]

I use two pre-registered measures of educational demand: the share of the eleven choices in which tutoring is selected and the continuous willingness-to-pay measure of Jack et al. (2022).

Comprehension. Before the tutoring task, respondents practiced the same choice mechanism using a pen and different quantities of marzipan candies. They then answered four questions about the binding row, the payment rule, the fact that choices could not be changed after the draw, and the one-in-ten payment probability.

Incorrect answers triggered the relevant instructions before the question was re-administered; the original answer remained recorded. The mean score was 3.8 out of 4.

Pre-registered hypotheses. The first pre-registered hypothesis is that Map-First lowers demand for tutoring. The second is that the effect is concentrated east of the boundary. I therefore estimate the pooled treatment effect and an interacted specification by side.

5.2.2 Experimental results

Figure 8 shows the treatment effects on the two pre-registered outcomes by side of the historical boundary. The first outcome is the share of multiple-price-list choices in which the respondent selects tutoring. The second is willingness to pay, estimated using the two-step procedure of Jack et al. (2022). Table E.1 in Appendix E reports the pooled specification, the interaction by side, and the Analco subsample.[24]

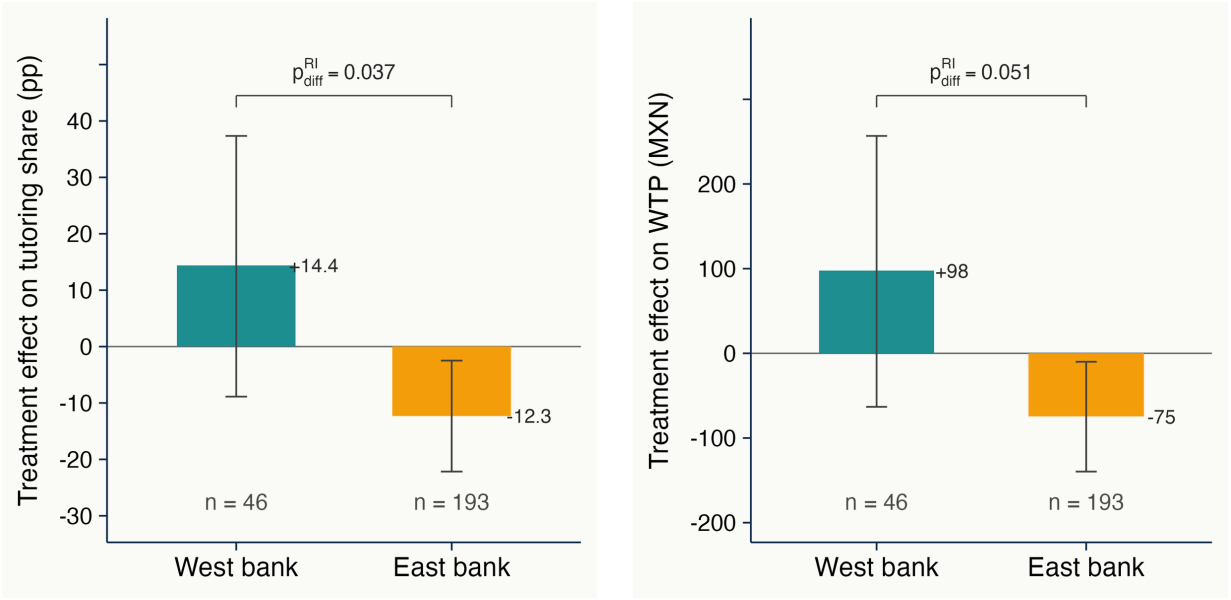
The regressions include enumerator and survey-day fixed effects, distance to the historical boundary, controls for the randomized presentation of the multiple-price list, and controls selected by LASSO on the outcome. Because treatment is randomized, these controls are included for precision.

Inference follows the assignment protocol. Map-First was randomized within side of the historical boundary, so I permute the treatment label within each bank, holding fixed the number treated on each side, and recompute the estimate 50,000 times. The reported p-values are the share of permutations producing an estimate at least as large in magnitude as the observed one. Table E.1 also reports conventional robust standard errors.

[23] Respondents drew a numbered slip from a sealed envelope to determine the binding row. The binding-row procedure and the one-in-ten payment probability were explained before the task. All households retained the MXN 100 participation payment regardless of the experimental outcome.

[24] I first recover each respondent's latent willingness to pay from a scaled random-effects probit of the eleven binary choices on gift-card value, estimated without treatment indicators. I then regress the resulting latent WTP on the treatment indicators. The reported standard errors bootstrap both steps at the respondent level. Randomization inference does not require re-estimating the first step because it contains no treatment indicators; under permutation only the second step changes. The bounded area-under-demand measure, used in earlier drafts, gives the same qualitative result (Table E.2).

Figure 8: Effects of neighborhood salience on educational demand



(a) Share of tutoring choices

(b) Willingness to pay

Notes: Treatment effects of the Map-First arm on the two pre-registered outcomes, by side of the historical boundary. The west-bank bar reports the Map-First coefficient; the east-bank bar reports Map-First + Map-First×Analco from the interacted specification in Table E.1. Panel (a) reports the share of the eleven multiple-price-list choices in which the respondent selected tutoring, in percentage points. Panel (b) reports latent willingness to pay in MXN using the two-step estimator of Jack et al. (2022). The regressions include enumerator and survey-day fixed effects, distance to the historical boundary, controls selected by LASSO on the outcome, and controls for the randomized presentation of the choice list. Inference is by randomization inference. Whiskers are 95 percent intervals obtained by inverting the permutation test, and the bracket reports the randomization-inference p-value for equality of the two side-specific treatment effects. The Map-First label is permuted 50,000 times within side of the historical boundary, holding the number treated on each bank fixed. Labels below the bars report the number of respondents on each side.

Treatment effects. Pooling both sides, Map-First lowers the tutoring share by 7.7 percentage points and willingness to pay by MXN 45, although both pooled estimates are imprecise.

On the east bank, Map-First lowers the tutoring share by 12.3 percentage points and willingness to pay by MXN 75. On the west bank, both estimates are positive and imprecise. Restricting the sample to the east bank gives similar estimates: the tutoring share falls by 11.7 percentage points and willingness to pay by MXN 71 (Table E.1).

Magnitude. Among east-bank respondents, the tutoring share falls from 0.573 in Map-After to 0.450 in Map-First, a decline of about one fifth. Latent willingness to pay falls from MXN 176 to MXN 102, or about 42%.

The effects correspond to 0.33 standard deviations for the tutoring share and 0.30 for willingness to pay, using Glass's Δ. The MXN 75 decline in willingness to pay is about one quarter of Mexico's 2026 general daily minimum wage.[25]

Interpretation. Random assignment to Map-First lowers educational demand on the historically excluded side. In the framework, this is the response produced when greater neighborhood salience places more weight on a positive neighborhood pull. The positive and imprecise west-bank estimates are also allowed by the model when the local pull is weak or negative.

The experimental coefficient is an intention-to-treat effect. It is not an estimate of Π_ℓ or $a(s)$, does not identify q_ℓ, and does not distinguish the local-network and conformity terms in (4). Nor does the experiment identify how the local state represented by q_ℓ arose historically.

5.2.3 What does the mapping task make salient?

Priming experiments can identify a behavioral response without isolating the precise content brought to mind (Cohn and Maréchal, 2016). I use respondents' polygons and their open-ended descriptions to characterize the object activated by the mapping task. These measures are descriptive because both treatment arms answer the follow-up questions immediately after drawing.

Respondents draw compact areas around their homes. The median polygon covers just over half a square kilometer, 88% contain the respondent's residence, and only 11% cross the historical boundary (Figure E.2 and Table E.3).

Their descriptions are similarly local. Using a pre-specified large-language-model classification following Haaland et al. (2025), 62% describe the neighborhood by its physical extent and 49% by everyday use; 13% mention people or belonging and 6% safety (Table E.4).[26]

[25] The general daily minimum wage effective January 1, 2026 was set by the *Comisión Nacional de los Salarios Mínimos* at MXN 315.04 and published in the *Diario Oficial de la Federación* on December 9, 2025. Guadalajara lies in the general minimum-wage zone rather than the northern-border free zone.

[26] I use `claude-opus-4-8`. An independent second classification yields mean Cohen's $\kappa = 0.93$ across themes;

Descriptions in terms of physical extent and everyday use are not evidence against a social object. In urban sociology, a neighborhood is constituted by routine interaction grounded in shared residence and the common use of local facilities, rather than by dense, intimate bonds (Smith, 2010; Sampson et al., 2002). The mapping task therefore centers attention on a compact area around the respondent's home. The prompts themselves do not mention the historical boundary, education, or a city-wide social identity.

5.2.4 Randomization, implementation, and robustness

Randomization balance. The Map-First and Map-After groups are balanced across more than thirty pre-treatment characteristics, including location, household composition, parental education, income, housing conditions, focal-child characteristics, and multiple-price-list comprehension. No individual difference remains statistically significant after false-discovery-rate adjustment, and treatment assignment is also orthogonal to enumerator (Tables E.5 and E.6).

Respondents answer, on average, 3.8 of four comprehension questions correctly, with no difference across arms. Never-switching and multiple-switching behavior is also balanced and close to the patterns reported by Jack et al. (2022) (Table E.6 and Figure E.3). There is no evidence that Map-First changes comprehension or use of the choice list.

Influential observations. Given the sample size, a few households could in principle drive the estimates. In a leave-one-out exercise, the estimated east-bank effect remains stable when any single household is removed (Table E.7).

Bounded willingness-to-pay measure. The MXN 250 upper limit of the multiple-price list compresses differences among respondents whose valuation lies at or above the top of the list. The latent estimator is not bounded by the grid, which is one reason it is the pre-registered primary measure. As a companion, I re-estimate the effect with the bounded area under each respondent's demand curve, which is also well defined for never- and multiple-switchers. The estimates are smaller, as the ceiling implies, but have the same sign, and the east-bank effect remains statistically significant (Table E.2).

Experimenter demand. Map-First respondents may infer that the study links neighborhoods to education and alter their choices accordingly. At the end of the survey, respondents were asked what they thought the study was about. I classify these responses using the same large-language-model procedure used for the neighborhood descriptions. Thirty of 239 respondents inferred a connection between neighborhood and education.

On the west bank, Map-First raises this inference by 22 percentage points. On the east bank, where the behavioral effect appears, the increase is essentially zero. The east-bank treatment

details are in Table E.4 and Section E.6.1.

effect is also essentially unchanged when I control for inferred study purpose or exclude these respondents (Table E.8). This pattern is difficult to reconcile with experimenter demand as the source of the east-bank effect.

5.3 Boundary salience and explanations of schooling

The tutoring experiment varies whether respondents have just thought about their neighborhood. A separate pre-registered exercise varies whether the historical divide is salient before respondents explain why a young person from their neighborhood might not continue studying after high school.

I randomize whether respondents are asked, before or after that open-ended question, how often they hear or use the expression "on the other side of the Calzada" (*de la Calzada para allá*), the expression associated with the historical divide described in Section 2. Nearly 80% report hearing or using it.

I classify the answers using a fixed codebook covering material constraints, the local place and social environment, the family, and the young person's own choices. On the Analco side, making the historical divide salient reduces references to money or the need to work by 15 percentage points, from 91 to 76%. References to the neighborhood rise from 8 to 14%, and references to family from 10 to 14%, although these increases are individually imprecise. Combining the neighborhood and economic categories, the within-response difference shifts by 21 percentage points toward neighborhood explanations. On the west bank, references to material constraints fall by 2 percentage points (Table E.9). Controls for the earlier map assignment, interview length, and respondent demographics leave the estimates similar.

The west-bank estimates are too imprecise to establish a differential response across sides. Within Analco, however, making the historical divide salient changes how respondents describe educational non-continuation: material explanations become markedly less frequent, and the balance between neighborhood and material explanations shifts toward the neighborhood. This exercise identifies an effect of boundary salience on explanations. It does not identify the historical assignment's effect on those explanations or any component of the neighborhood pull.

5.4 The local social environment around the boundary

The experiment establishes that mapping the neighborhood immediately before the choice reduces educational demand on the east bank. It does not identify what feature of the neighborhood generates the response. The remaining survey modules compare the two sides on local exchange, educational beliefs, stated norms, and attachment. These comparisons are observational and are not mediators of the randomized treatment.

Locally oriented giving. Following Enke et al. (2023), respondents complete three hypothetical allocation exercises. In the first, they divide MXN 500 between two equally well-off strangers, one from their own *barrio* and one from elsewhere in Guadalajara. In the second, they divide MXN 500 between a Guadalajara resident and someone from outside the city. The order of these two exercises is randomized. A third question asks how much of an unexpected MXN 500 windfall the respondent would donate to charity.

Panel A of Table 3 takes the three allocation questions in turn. Analco residents donate no less to charity, so what follows is not a difference in generalized altruism. They allocate about MXN 20 more to the stranger from their own *barrio*. The equal split is modal on both sides, and the departure from it is where the sides separate most sharply: no west-bank respondent gives the *barrio* stranger strictly more than half, compared with 17% of Analco respondents. At the Guadalajara radius the side difference is positive but imprecisely estimated, so the exercise does not establish that parochialism stops at the *barrio*. Measured relative to what each respondent allocates to the city stranger, however, the tilt toward the *barrio* is about nine percentage points larger on the excluded side. The panel therefore shows a difference in how prosociality is tilted toward the nearer party rather than in how much is given away.

Local exchange. Panel B reports direct measures of interaction with neighbors. Analco respondents report receiving help from neighbors about twice as often over the previous year and spending close to an additional hour per week with them. These measures provide a more local view of exchange near the historical boundary than the municipality-wide survey evidence in Section 4.

Two vignettes ask about a family whose child takes a job near home and a family whose child leaves for university elsewhere in the city. Respondents report which family neighbors would help first and which decision would strain neighborly relations more. Respondents on average place the stayer's family ahead on help and the university decision ahead on strain, but neither answer differs precisely across the boundary. The clearer side differences are in realized exchange: help received from neighbors and time spent with them.

Aspirations and expectations. Panel C shows little difference in parental aspirations. For a randomly selected focal child, parents on both sides report similarly high aspirations and expectations for upper-secondary completion and university. The estimates are small and statistically insignificant.

Perceived returns and the university threshold. Panel D provides little evidence that Analco parents perceive lower monetary returns to education. Expected earnings at age 30 for high-school and university graduates, the implied university premium, and beliefs about whether an otherwise identical Analco graduate earns less than a western-side graduate are statistically similar across

Table 3: Localized prosociality and alternative mechanisms

	Mean west bank	Mean Analco	Analco coefficient	q
Panel A. Localized prosociality (allocation exercises)				
Donation to a good cause	344.6	371.8	23.0 (24.3)	0.16
Allocation to a barrio stranger	244.6	262.9	19.7** (9.0)	0.06
Gives strictly more to the barrio stranger	0.000	0.166	0.170*** (0.032)	0.00
Allocation to a city stranger	251.1	265.6	14.6 (12.0)	0.13
Relative barrio premium	-0.015	0.053	0.088** (0.044)	0.07
Panel B. Localized network benefits and the cost of leaving				
Times helped by neighbors, past 12 months	0.97	1.96	1.10** (0.44)	0.05
Hours per week with neighbors	0.90	1.66	0.88* (0.47)	0.11
Vignette: the stayer's family gets more help	0.52	0.60	0.17 (0.23)	0.35
Vignette: leaving for university strains relations	0.59	0.72	0.16 (0.25)	0.35
Panel C. Aspirations and expectations				
Aspires to high school or more	1.000	0.990	−0.007 (0.007)	1.00
Aspires to university	0.935	0.948	0.027 (0.036)	1.00
Expects high school or more	0.957	0.953	0.010 (0.027)	1.00
Expects university	0.870	0.850	0.020 (0.049)	1.00
Panel D. Returns to schooling and investment threshold				
Expected earnings, high-school graduate†	12,924	12,690	−0.045 (0.081)	1.00
Expected earnings, university graduate†	24,174	24,503	0.021 (0.078)	1.00
Perceived return to university (uni. − HS)	11,250	11,812	1, 051 (1,657)	1.00
Perceives excluded-side graduate earns less	0.20	0.23	0.091 (0.177)	1.00
Reservation salary to justify university†	31,043	36,363	0.166** (0.071)	0.10
Panel E. Perceived neighbor norms (pluralistic ignorance)				
Personal norm: youth should work, not study (1–7)	2.61	2.55	−0.237 (0.311)	1.00
Perceived neighbors sharing that norm (0–10)	5.57	5.70	0.163 (0.355)	1.00
Pluralistic ignorance (perceived − own, both on 0–10)	2.88	3.11	0.558 (0.636)	1.00
Panel F. Stated attachment to the neighborhood				
Feels part of this colonia (1–7)	5.76	5.72	−0.137 (0.253)	1.00
Would prefer to live in another colonia (1–7)	3.22	3.16	−0.077 (0.437)	1.00
Makes an effort so children know the neighbors (1–7)	4.48	4.92	0.258 (0.377)	1.00

Notes: The first two columns report unadjusted means for west-bank and Analco respondents. The third column reports the coefficient on the Analco indicator, controlling for Map-First assignment, respondent sex, age and age squared, randomized question versions where applicable, and respondent-education fixed effects. Panel C additionally controls for the focal child's sex, age, and current school attendance. In Panel E the personal norm is elicited on a 1–7 agreement scale and the perceived neighbor norm as a count out of ten; the personal norm is linearly rescaled to 0–10 before the difference is taken, so both enter the pluralistic-ignorance row on a common metric. Because the two questions use different response formats, the level of that difference depends on the mapping between them; the interpretable quantity is the comparison across banks, not the size of the gap itself. HC1 robust standard errors are in parentheses. The estimation sample contains 238 households. Outcomes marked † are estimated by Poisson pseudo-maximum-likelihood, so coefficients are semi-elasticities while the reported means are in MXN; all other outcomes are estimated by OLS in levels. The q column reports sharpened false-discovery-rate values (Anderson, 2008), computed separately within each panel. All estimates are observational cross-boundary comparisons. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

sides.

Parents on both sides also report returns well above the realized cross-sectional premium. The median respondent expects a university graduate to earn 100% more than a high-school graduate at age 30, compared with a 57% premium in 2020 census microdata for full-time Guadalajara workers aged 25–35 (Figure A.10 in Appendix A).

One measure differs. Analco respondents require an approximately 18% higher salary at age 30 before they consider university worthwhile.[27] Because perceived monetary returns are otherwise similar, the higher reservation threshold is consistent with an additional nonpecuniary hurdle to university investment. The question does not identify that hurdle or connect it structurally to Π_ℓ.

Perceived educational norms. Following Bursztyn and Yang (2022), respondents report both their own and their perceived neighbors' agreement with the statement that young people should work rather than continue studying. Panel E shows no systematic side differences. Analco residents are no more likely to endorse work-over-study norms, do not perceive their neighbors as less supportive of education, and do not display greater pluralistic ignorance.

These questions measure stated beliefs about explicit work-versus-study norms. They do not measure the local educational behavior $\bar{e}_\ell$ in (4).

Stated attachment. Panel F measures attachment to the *colonia*. Respondents on both sides report similar feelings of belonging, similar preferences over living elsewhere, and similar effort to have their children know their neighbors. The differences in local exchange and locally oriented giving therefore do not appear as a corresponding difference in stated attachment.

Scale of the university hurdle. The reservation-salary difference is about MXN 5,300 per month, roughly 45% of the university wage premium respondents expect. The tutoring experiment concerns a much smaller and shorter-term investment, so the two quantities are not directly comparable. For reference, Map-First lowers WTP for tutoring by about 42% on the east bank.

A second calculation asks how large the stated university threshold is relative to the schooling discontinuity. Among Analco respondents, 18% expect university earnings that clear their own reported threshold. Subtracting the MXN 5,300 cross-bank difference in mean reservation salary from each of those thresholds raises the share to 37%. If every respondent who newly clears the threshold completed the four years a degree adds beyond high school, the 19-percentage-point difference would amount to about 0.75 years of schooling per Analco resident, close to the 0.8-year discontinuity at the historical boundary.

[27] The question describes a local young person who has completed high school and received a job offer but could instead attend university. The prompt states that university entails tuition, materials, transportation, and forgone earnings. Respondents are then asked how much the young person would need to earn per month at age 30 for university to have been worthwhile.

The two numbers are not estimates of a common quantity. The calculation converts a one-time crossing of a stated threshold into completed years, while the boundary estimate averages realized schooling across many margins among residents aged 6 and older. I treat the proximity as illustrative of scale rather than as a test. The reservation salary is self-reported, the exercise concerns one educational margin, and the calculation does not identify how much of the RDD effect operates through local social considerations.

Summary. The clearest side differences in these modules concern local exchange. Analco residents receive more help from neighbors, spend more time with them, and direct more giving toward a stranger from their own neighborhood. They do not report lower educational aspirations, lower perceived monetary returns to schooling, stronger explicit work-versus-study norms, or stronger stated attachment. The modules characterize the local environment around the boundary; they do not identify which feature generates the experimental response.

5.5 Persistence

The static framework explains why neighborhood salience can affect educational demand today. The dynamic part of the model asks how a local state capable of generating that response can coexist with the disappearance of the original provision gap.

Historical provision and education. The starting point is the historical asymmetry in formal educational institutions documented in Section 4.1. That asymmetry was one element of a difference in colonial standing, and the model isolates it: poorer provision raises the effective cost $C(P_\ell)$ of education. This can create an initial education difference even when the two neighborhoods begin with the same composition,

$$q_A^0 = q_B^0 = \bar{q}.$$

An earlier period in which schooling is very costly or infeasible can leave a low schooling norm. Once educational investment becomes feasible but provision remains weaker, the eastern side can remain on the low-education branch while the western side is on the high branch. This argument does not require the historical record to contain literally zero schools on the eastern side; it requires a sufficiently large difference in the cost and feasibility of educational investment.

Transmission and feedback. The appendix introduces intergenerational transmission using Bisin and Verdier (2001). In the low-education configuration, the incentive to transmit neighborhood orientation is stronger because an oriented child retains more of the locally valuable payoff and does not depart from the local education norm.

This does not imply that orientation rises whenever education is low. The low branch contains a tipping point q_c. When the local stock exceeds that point, orientation can accumulate toward a

stable rest point; below it, orientation declines even though the low-education equilibrium remains available. The feedback is therefore conditional: education affects orientation ($e \longrightarrow q$) by changing transmission incentives, while orientation affects education ($q \longrightarrow e$) by strengthening the neighborhood pull. The empirical analysis does not identify either direction separately.

Provision convergence and residential churn. When provision converges, the cost wedge represented by $C(P_A) - C(P_B)$ disappears. The local state need not disappear at the same time. If the accumulated q_A remains high enough to support the low branch under common provision, different educational configurations can remain.

Residential turnover works against that persistence. Let μ_ℓ be the fraction of residents replaced each period by draws from the city distribution with neighborhood-oriented share $\bar{q}$. Then

$$q_{\ell,t+1} - q_{\ell,t} = (1 - \mu_\ell)\Phi_r(q_{\ell,t}) + \mu_\ell(\bar{q} - q_{\ell,t}), \tag{7}$$

where Φ_r is the regime-specific transmission flow.

On the low branch, transmission can regenerate the local state while churn pulls it toward the city average. Greater churn lowers any stable persistent rest point. When turnover is sufficiently large, the persistent configuration disappears and q_ℓ moves toward $\bar{q}$. On the high branch, the minimal model has zero endogenous transmission drift because both types make the same educational choice; composition can still change through the churn term. The appendix derives these conditions and discusses the simplifying nature of the zero high-branch drift.

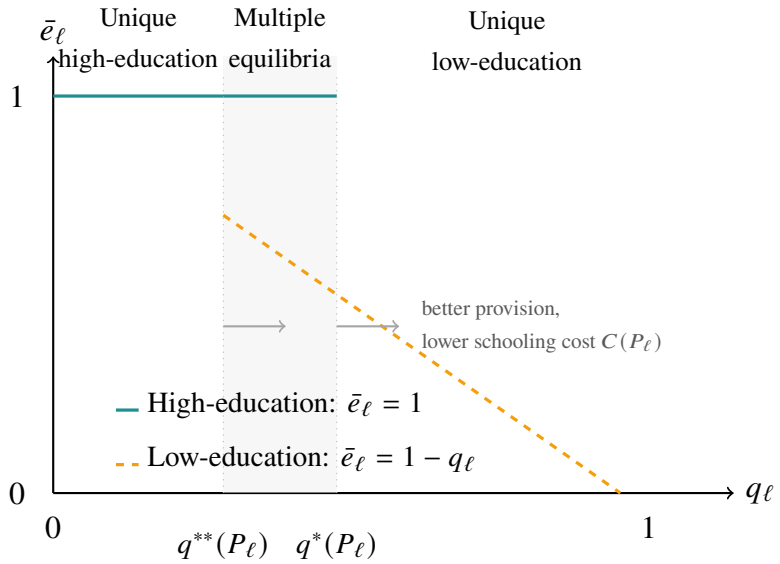
Persistence in the data. Two empirical patterns are consistent with slow local turnover. First, a proxy for long-term residence is higher east of the historical boundary (Table B.8), although Appendix B.12 explains why that proxy cannot be read as a measured difference in turnover. In the model, lower turnover weakens the force pulling local composition toward the city distribution. Second, paternal surnames remain spatially distinctive well into the twentieth century (Figure 7), consistent with persistent residential lineages across the divide.

Other relational markers need not move at the same rate. Given-name distinctiveness fades after 1880 (Section 4.2), while paternal surnames remain spatially distinct and present-day surveys document other forms of local relational difference. The model does not treat these outcomes as repeated observations of the same state variable.

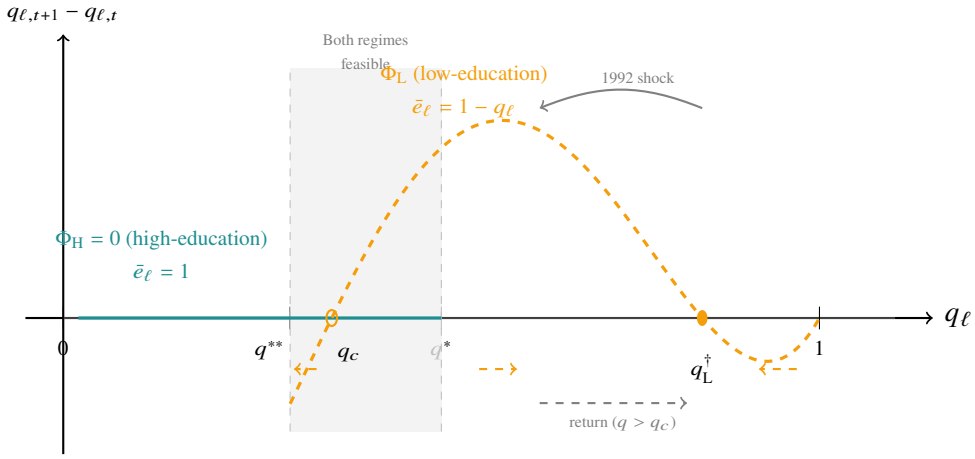
The 1992 Reforma explosions. In April 1992, a series of explosions along the Reforma corridor in eastern Guadalajara destroyed roughly 1,500 buildings across several kilometers of Analco. The rebuilt area was subsequently occupied largely by new residents rather than by the displaced families (Reguillo, 1998). Panel (a) of Figure 10 shows the affected area.

I compare individuals in the directly affected blocks with individuals in their second- through fifth-degree neighboring blocks, all within Analco and within two kilometers of the historical

Figure 9: Educational regimes and neighborhood-orientation dynamics



(a) Educational regimes



(b) Dynamics of the neighborhood-oriented share

Notes: Panel (a) plots equilibrium education $\bar{e}_\ell$ against the share of neighborhood-oriented residents q_ℓ. When q_ℓ is low, only the high-education regime exists; when q_ℓ is high, only the low-education regime exists. In the shaded region $[q^{**}, q^*]$, both are feasible. Better public provision shifts both thresholds to the right because it lowers the effective cost of educational investment, $C'(P) < 0$. Panel (b) plots the endogenous transmission flow at zero residential churn. On the low branch, orientation rises only for $q_\ell \in (q_c, q^\dagger)$; below q_c it falls even though the low-education regime may still exist. In the minimal high-education branch, both types make the same educational choice, so endogenous transmission drift is zero. Residential churn, omitted from the panel, continues to move composition toward the city average. Parameters are $\rho - C(P) = 0.2$, $\kappa = 0.4$, $\delta = 0.6$, and $\Gamma = 2$. They are illustrative, not calibrated. Derivations are in Appendix D.

boundary; first-degree neighbors are excluded as a buffer. The regressions pool the 2000, 2010, and 2020 censuses at the block level, include year, age, and sex fixed effects, and cluster standard errors by block. The 1990 census provides a pre-shock comparison, but only at the census-tract (AGEB) level: it contrasts the tracts the corridor would later cross with the remaining Analco tracts. That partition is coarser than, and not nested in, the block-level contrast used from 2000 onward, so it speaks to whether the broader areas differed before the explosions rather than to balance between the block-level groups compared afterward.

At that coarser resolution there is no pre-shock schooling difference between the tracts the corridor would cross and the rest of Analco in 1990. By 2000, adults in the rebuilt blocks have almost one additional year of schooling. The difference disappears by 2010 and 2020.

Children show a similar pattern. Conditional on parental schooling, attendance in the rebuilt blocks is about five percentage points higher in 2000 than in the comparison blocks. That difference also fades in later census waves.

The episode is consistent with a temporary change in residential composition followed by convergence toward the surrounding area. I do not interpret it as a clean test of the churn mechanism. The explosions combined destruction, displacement, compensation, reconstruction, housing changes, and repopulation, and the census follows places rather than displaced households. This differs from settings such as Chyn (2018) and Nakamura et al. (2022), which follow displaced individuals.

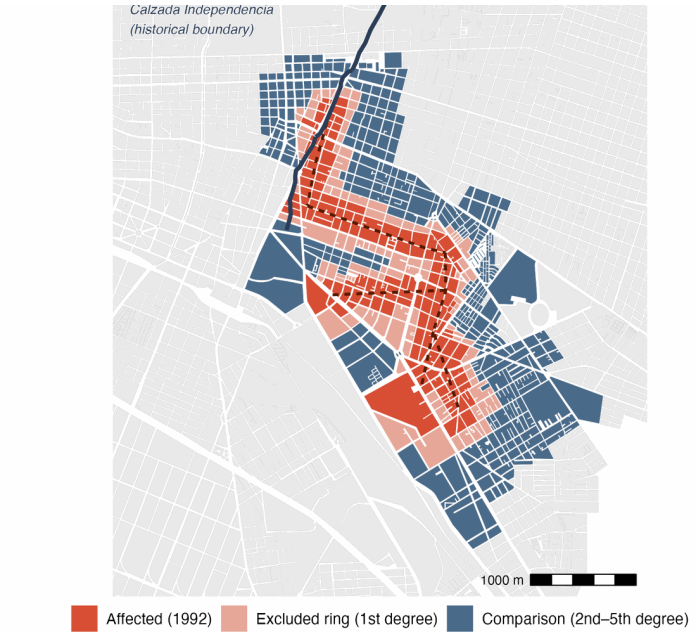
The narrower result is that rebuilt blocks were more positively selected on adult schooling by 2000 and had higher child attendance conditional on parental schooling, and that both differences faded. Mean reversion and a housing cycle in which new construction first attracts more educated residents and later loses them remain alternative interpretations.

6 Neighborhood Orientation Beyond Guadalajara

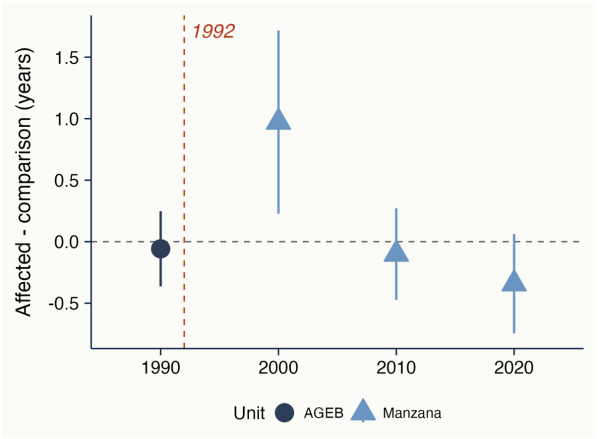
The experiment identifies how neighborhood salience affects educational demand in one setting shaped by historical exclusion. I next ask whether neighborhood orientation is also associated with lower schooling elsewhere. I then examine whether this association is stronger under residential segregation, a setting in which exclusion can operate both through unequal access to public goods and through limited social mixing. These exercises are descriptive, but they extend the two margins emphasized in Guadalajara beyond its particular historical setting. I study subnational regions worldwide, Mexican municipalities, U.S. ZIP codes, and individuals.

For the place-level analysis, I use the Facebook Social Connectedness Index (Bailey et al., 2018), which measures the relative probability that two people in a pair of locations are friends. For each place, I calculate the share of friendship ties contained within a short radius. This captures the extent to which a place's social ties are concentrated nearby, one observable dimension of neighborhood orientation. I relate this measure to schooling within larger geographic units, so the comparisons are between places exposed to the same broader national or regional environment.

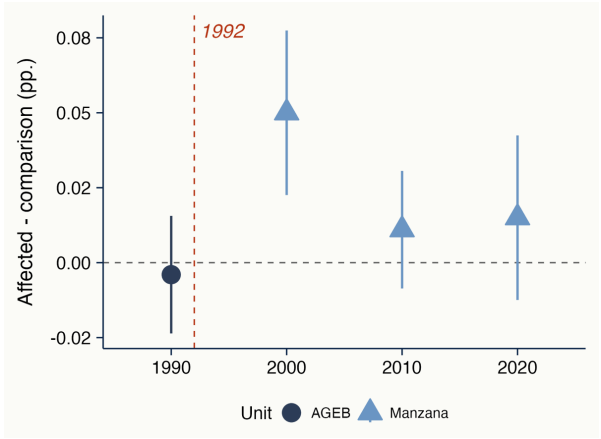
Figure 10: The 1992 Reforma explosions: affected blocks, selection, and school attendance



(a) Affected and comparison blocks



(b) Selection: years of schooling



(c) School attendance of children

Notes: Panel (a) maps the blocks directly affected by the April 1992 explosions (red), defined as any intersection with an 80-meter buffer around the collector's path (dashed line); the immediate first-degree neighbors, excluded from the comparison (yellow); and the second- through fifth-degree neighboring blocks used as the comparison group (teal). The dark line is Calzada Independencia, the historical boundary; the affected corridor lies entirely on the eastern (Analco) side. Panels (b) and (c) plot the difference between affected and comparison blocks by census wave: panel (b) reports years of schooling among adults aged 18 and older, and panel (c) reports school attendance among children aged 6–18, conditional on parental years of schooling. The 1990 estimate is a pre-shock placebo at the AGEB level; the 2000–2020 estimates are at the block (*manzana*) level. Vertical bars show 95 percent confidence intervals; the dashed vertical line marks the explosions. Design details are in Appendix B.13.

Localization of social ties and schooling. The relationship is negative across spatial scales. Across subnational regions worldwide, comparing regions within the same country, a greater concentration of friendship ties nearby is associated with fewer years of schooling (Figure 11a and Table F.1). Because these schooling data cover primarily developing and middle-income countries, I repeat the exercise across European NUTS2 regions, where tertiary attainment also declines as social ties become more localized (Table F.2).

The same pattern appears within Mexico. Across roughly 2,200 municipalities, a larger share of friendship ties contained within the municipality is associated with about 1.5 fewer years of schooling (Table F.3). These estimates are observational, but the negative association between local orientation and schooling is not specific to Guadalajara.

Residential segregation. I next ask whether this relationship is stronger where residential segregation is higher. Segregation is useful here because it can operate along the same two margins emphasized in Guadalajara. It can limit social mixing across neighborhoods and allow local differences in social interaction to persist. It can also produce unequal access to public goods and other neighborhood resources (Chyn et al., 2026). In this sense, segregation provides a contemporary setting in which exclusion can remain both material and social. Segregation has also been associated with worse institutional outcomes across countries (Alesina and Zhuravskaya, 2011).

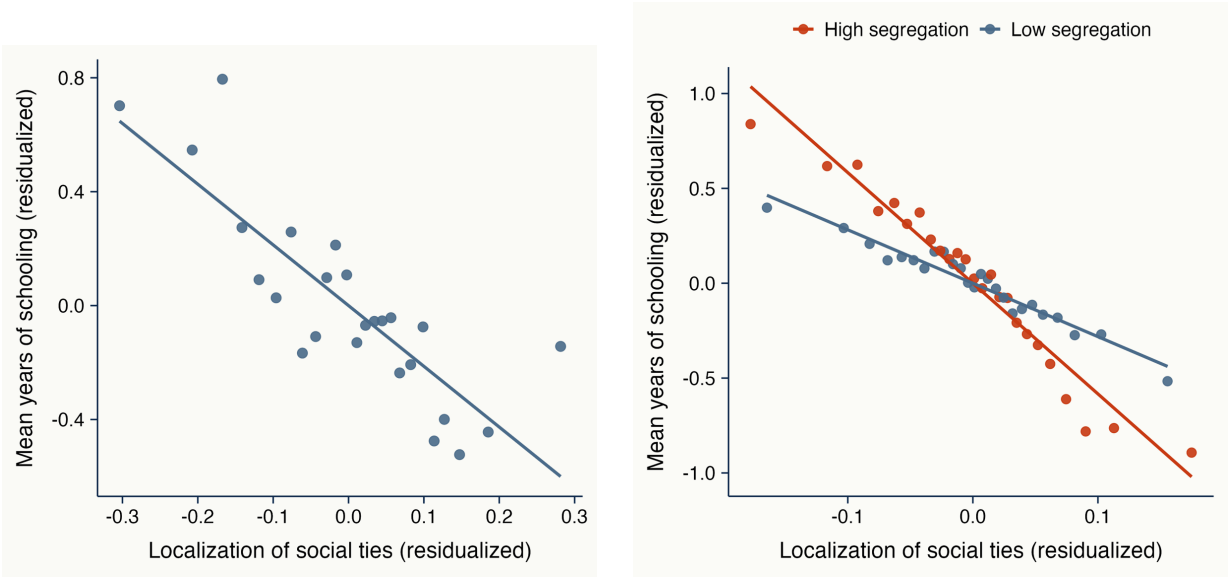
Across U.S. ZIP codes, comparing ZIPs within the same county, the negative relationship between schooling and the localization of social ties is roughly twice as steep in more segregated counties. The same pattern appears using either racial or income segregation (Figure 11b and Table F.4).

Individual evidence. I examine the same relationship at the individual level using Wave 7 of the World Values Survey. Forty-two countries have the variables the exercise requires; the ethnic-group and parental-education fixed effects absorb seven of them entirely, because the survey does not record those variables there, leaving thirty-five countries in the estimates (Appendix F.1). I construct a *neighborhood trust premium*, defined as standardized trust in neighbors minus average trust in family and personally known individuals. Unlike the Facebook measure, this captures the relative orientation of trust toward neighbors rather than the geographic concentration of friendships. It closely parallels the neighborhood trust premium measured in Guadalajara in Section 4.3.

The outcomes are years of schooling, educational mobility relative to parents, and dropout before upper secondary. I classify countries above the mean of the Alesina and Zhuravskaya (2011) ethnic-segregation index as high segregation. The specifications include a rich set of fixed effects and control flexibly for mother's and father's education. They exclude respondent occupation, which is realized after schooling.

In the pooled sample, the neighborhood trust premium is negatively associated with educational

Figure 11: Localization of social ties and schooling



(a) Subnational regions, within country

(b) U.S. ZIP codes, by county segregation

Notes: Binned scatters of mean years of schooling against the localization of a place's Facebook friendship network (Bailey et al., 2018), with both axes residualized on fixed effects and controls. Panel (a) compares subnational regions within country, controlling for log population and log area; schooling is from the Global Data Lab. Panel (b) compares U.S. ZIP codes within county, with the same controls and schooling from the ACS. ZIP codes are shown separately for counties in the bottom and top terciles of the white/non-white residential-dissimilarity index. The schooling gradient is negative in both groups and roughly twice as steep in high-segregation counties. Full regressions are in Tables F.1 and F.4.

Table 4: Global evidence: the neighborhood trust premium and human capital accumulation

	Years of Education				Upward Educational Mobility				Dropout Before Upper Secondary			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Neighborhood trust premium (vs. close ties)	-0.018***	-0.025***	-0.022***		-0.020***	-0.026***	-0.022***		0.014***	0.019***	0.014***	
	(0.004)	(0.004)	(0.004)		(0.005)	(0.005)	(0.005)		(0.004)	(0.004)	(0.004)	
Out-group trust		0.051***	0.047***			0.044***	0.039***			-0.041***	-0.034***	
		(0.004)	(0.005)			(0.005)	(0.006)			(0.005)	(0.005)	
Trust in close ties (family + known)			0.011**				0.014***				-0.018***	
			(0.005)				(0.006)				(0.005)	
Neighborhood trust premium × Low ethnic segregation				-0.011**				-0.014**				0.005
				(0.005)				(0.006)				(0.005)
Neighborhood trust premium × High ethnic segregation				-0.038***				-0.034***				0.027***
				(0.007)				(0.008)				(0.008)
Observations	47,072	46,952	46,952	46,952	47,054	46,934	46,934	46,934	47,072	46,952	46,952	46,952
R^2	0.5072	0.5092	0.5093	0.5095	0.3145	0.3160	0.3162	0.3163	0.4355	0.4369	0.4371	0.4372
Within R^2	0.0006	0.0042	0.0044	0.0049	0.0005	0.0024	0.0026	0.0028	0.0003	0.0023	0.0027	0.0029
Subnational Region-Survey year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Subnational Region-Settlement size fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gender-Age fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Country of birth fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ethnic group fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Religious group fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Father's education-Mother's education fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: OLS estimates from the World Values Survey / European Values Study integrated time series, wave 7 (2017–2023). The unit of observation is the respondent, and all outcomes are standardized. The neighborhood trust premium is standardized trust in neighbors minus average trust in family and in people known personally, so it measures the relative orientation of trust toward neighbors rather than the level of trust. A country is classified as high ethnic segregation if its Alesina and Zhuravskaya (2011) ethnic-segregation index is at or above the cross-country mean, and as low ethnic segregation otherwise. Each outcome block reports a stepwise sequence — (1) the neighborhood trust premium alone, (2) adding out-group trust, (3) adding trust in close ties — and (4) a specification in which all three trust terms are interacted with the low- and high-segregation indicators. Column (4) therefore reports separate slopes by segregation regime rather than a base category and a differential; the displayed rows give the low- and high-segregation slopes of the neighborhood trust premium, while the corresponding interactions for out-group trust and for trust in close ties are included in the regression and omitted from the displayed rows. All specifications include region-by-year, region-by-settlement-size, sex-by-age, country-of-birth, ethnicity, religion, and father-by-mother-education fixed effects. Observations are weighted by survey weights. Heteroskedasticity-robust standard errors, without clustering, are reported below each coefficient in parentheses.

attainment (Table 4). This average masks substantial heterogeneity by segregation. A one-standard-deviation increase in the neighborhood trust premium is associated with 0.038 standard deviations fewer years of schooling in high-segregation countries, against 0.011 in low-segregation ones, and with 0.034 against 0.014 of educational mobility. Dropout rises by 0.027 standard deviations in high-segregation countries and is unrelated to the premium elsewhere. Table F.6 shows that the contrast between the two regimes survives dropping the fixed effects that cost the sample seven countries.

Both levels of trust are associated with better outcomes. Out-group trust predicts higher attainment, and trust in close ties does too, with a smaller coefficient in each of the three outcomes. The negative association is confined to the relative orientation of trust toward neighbors. The larger out-group coefficient is related to Chetty et al. (2022), who find that upward mobility is associated with connectedness to higher-SES contacts rather than the cohesion of close networks.

Interpretation. These exercises are descriptive. The Facebook measure captures where social ties are concentrated, while the World Values Survey captures the relative orientation of trust toward neighbors. Neither is a direct measure of the theoretical state q, and residential segregation is not an exogenous source of variation in neighborhood orientation. The estimates therefore do not show that segregation produces more localized social ties or that either produces lower schooling.

They do show two broader patterns. Across several levels of aggregation, schooling is lower where social ties are more locally concentrated. In the U.S. ZIP-code and individual exercises, the negative relationship is stronger in more segregated settings. Segregation is relevant to the comparison because it can combine unequal access to public goods and neighborhood resources with more limited social mixing, the two dimensions of exclusion emphasized in Guadalajara.

The role of the Guadalajara evidence is different. The historical design identifies the long-run reduced-form effect of the colonial assignment under the continuity assumption, while the experiment identifies the effect of randomized neighborhood salience on educational demand. The evidence in this section does not identify those links outside Guadalajara. It shows instead that the descriptive relationship between neighborhood orientation and schooling appears across other settings, and that it is more pronounced where residential segregation is higher.

7 Conclusion

This paper shows that educational gaps can persist long after differences in public provision disappear, and that neighborhood salience can reduce educational demand on the historically excluded side of a city. In Guadalajara, the colonial boundary still marks lower schooling and attendance nearly five centuries later, even though the two banks now share the same schooling system and measured school supply and quality no longer differ at the boundary. Historical

records and present-day surveys document relational differences across the same divide. In the incentivized choice experiment, making the neighborhood salient lowers educational demand on the excluded side, while a separate boundary-salience exercise reduces references to material constraints in explanations of educational non-continuation and shifts the balance of explanations toward the neighborhood.

The results matter for place-based policy. Greater social integration can improve children's outcomes (Chetty et al., 2026a), but in highly segregated settings substantial mixing may be slow, costly, or difficult to achieve at scale. The first-order response remains to equalize public provision. A second-order response may be to provide institutions that perform some of the functions otherwise supplied by the neighborhood pool, reducing the need for households to trade educational investment against locally valuable ties.

The findings also suggest caution around policies that formally label neighborhoods. Denmark's ghetto list and Sweden's police lists of vulnerable areas designate specific places as social problems, with documented effects on housing prices, residential sorting, and student outcomes (Andersson et al., 2026; Domínguez et al., 2025; Govind et al., 2025). Guadalajara does not identify the long-run effect of labeling itself, but the boundary-salience exercise shows that making a historically meaningful spatial designation salient can shift how residents understand educational disadvantage. This raises the possibility that place labels operate not only through markets and sorting, but also through the meanings attached to neighborhoods.

The experiment identifies the effect of neighborhood salience on educational demand, not which feature of local social life generates that response. The narrower implication is that material convergence need not eliminate the social margin through which neighborhoods continue to shape educational investment.

References

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson, "Do Immigrants Assimilate More Slowly Today Than in the Past?," *American Economic Review: Insights*, 2020, *2* (1), 125–141. (cited on p. 24).

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 12 2001, *91* (5), 1369–1401. (cited on pp. 5 and 10).

_ , _ , and _ , "Institutions as a Fundamental Cause of Long-Run Growth," 2005. (cited on p. 10).

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf, "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189. (cited on p. 10).

Akerlof, George A. and Rachel E. Kranton, "Economics and identity," *Quarterly Journal of Economics*, 2000, *115* (3), 715–753. (cited on p. 6).

Akerlof, George A and Rachel E Kranton, "Identity and Schooling: Some Lessons for the Economics of Education," *Journal of Economic Literature*, 12 2002, *40* (4), 1167–1201. (cited on p. 6).

Alberro, Solange, *Movilidad social y sociedades indígenas de Nueva España: Las élites, siglos XVI–XVIII*, Mexico City: El Colegio de México, 2019. (cited on p. 9).

Aldeco, Lucía, Mariana Calderón, Daniel Chiquiar, Gordon Hanson, Jorge Pérez Pérez, and Carlos Velázquez, "Local Labor Markets in Mexico: Definition, Databases, and Analysis," Working Paper, Banco de México, Mexico City 2024. (cited on pp. 7 and 19).

Alesina, Alberto and Ekaterina Zhuravskaya, "Segregation and the Quality of Government in a Cross Section of Countries," *American Economic Review*, August 2011, *101* (5), 1872–1911. (cited on p. 45).

Algan, Yann and Pierre Cahuc, "Trust, Growth, and Well-Being: New Evidence and Policy Implications," *Handbook of Economic Growth*, 1 2014, *2*, 49–120. (cited on p. 6).

Aliprantis, Dionissi, "Making Our Neighborhoods, Making Our Selves: A Review Essay," *Journal of Economic Literature*, 2023, *61* (3), 1172–1187. (cited on p. 5).

Ambrus, Attila, Erica Field, and Robert Gonzalez, "Loss in the Time of Cholera: Long-Run Impact of a Disease Epidemic on the Urban Landscape," *American Economic Review*, 2020, *110* (2), 475–525. (cited on p. 5).

Ananat, Elizabeth Oltmans, "The wrong side(s) of the tracks: The causal effects of racial segregation on urban poverty and inequality," *American Economic Journal: Applied Economics*, 4 2011, *3* (2), 34–66. (cited on p. 10).

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 12 2008, *103* (484), 1481–1495. (cited on pp. 16 and 26).

Anderson, Rodney D., "Race and Social Stratification: A Comparison of Working-Class Spaniards, Indians, and Castas in Guadalajara, Mexico in 1821," *The Hispanic American Historical Review*, 5 1988, *68* (2), 209. (cited on p. 7).

_ **and Tamara Spike**, "Making History Count: The Guadalajara Census Project (1791–1930)," *Hispanic American Historical Review*, 5 2007, *87* (2), 327–351. (cited on pp. 19, 24, and 25).

Andersson, Henrik, Ina Blind, Fabian Brunåker, Matz Dahlberg, Greta Fredriksson, Jakob Granath, and Che-Yuan Liang, "What Is in a Label? On Neighborhood Labeling, Stigma and Housing Prices," *Regional Science and Urban Economics*, 2026, *118*, 104191. (cited on p. 48).

Arias, Luz Marina and Desha M. Dirod, "Indigenous Origins of Colonial Institutions," *Quarterly Journal of Political Science*, 9 2014, *9* (3), 371–406. (cited on p. 7).

Atkin, David, Eve Colson-Sihra, and Moses Shayo, "How do we choose our identity? A revealed preference approach using food consumption," *Journal of Political Economy*, 2021, *129* (4), 1193–1251. (cited on p. 6).

Bagagli, Sara, "The (Express)Way to Segregation: Evidence from Chicag," Technical Report 2025. (cited on p. 10).

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, "Social Connectedness: Measurement, Determinants, and Applications in Economic Research," *Journal of Economic Perspectives*, 2018, *32* (3), 259–280. (cited on pp. 43 and 46).

Baldomero-Quintana, Luis, L. Guillermo Woo-Mora, and Enrique De la Rosa-Ramos, "Infrastructures of race? Colonial indigenous segregation and contemporary land values," *Regional Science and Urban Economics*, 1 2025, *110*, 104065. (cited on pp. 5 and 14).

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse, "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States," *Econometrica*, 11 2020, *88* (6), 2329–2368. (cited on p. 24).

Becker, Sascha O, P David Boll, and Hans-Joachim Voth, "Spatial Unit Roots in Regressions: A Practitioner's Guide and a Stata Package," 2025. (cited on p. 22).

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer, "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice," *American Economic Review*, 5 2024, *114* (5), 1281–1337. (cited on p. 5).

Bisin, Alberto and Thierry Verdier, ""Beyond the melting pot": Cultural transmission, marriage, and the evolution of ethnic and religious traits," *Quarterly Journal of Economics*, 8 2000, *115* (3), 955–988. (cited on p. 4).

_ **and** _ , "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, 4 2001, *97* (2), 298–319. (cited on pp. 4 and 40).

_ **and** _ , "Economic Models of Identity Formation and Cultural Transmission," 2026. Working paper, July 2026. (cited on p. 4).

Bleakley, Hoyt and Jeffrey Lin, "Portage and Path Dependence," *The Quarterly Journal of Economics*, 5 2012, *127* (2), 587–644. (cited on pp. 5 and 9).

Borusyak, Kirill and Peter Hull, "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 11 2023, *91* (6), 2155–2185. (cited on p. 22).

Bretón, Jorge Alberto Trujillo, "La ciudad del pecado: los bajos fondos y las drogas enervantes en Guadalajara, 1915–1946," in Jorge Alberto Trujillo Bretón, ed., *Por el mundo del delito y sus pormenores: historia, marginalidad y delito en América Latina*, Guadalajara, Mexico: Universidad de Guadalajara, 2018, pp. 307–336. (cited on p. 10).

Buringh, E., "Urbanisation hub: The clio-infra database on urban settlement sizes: 1500–2000.," 2018. (cited on p. 9).

Bursztyn, Leonardo and David Y. Yang, "Misperceptions About Others," *Annual Review of Economics*, 8 2022, *14*, 425–452. (cited on p. 39).

_ **and Robert Jensen**, "How Does Peer Pressure Affect Educational Investments?," *The Quarterly Journal of Economics*, 2015, *130* (3), 1329–1367. (cited on p. 5).

_ **, Georgy Egorov, and Robert Jensen**, "Cool to Be Smart or Smart to Be Cool? Understanding Peer Pressure in Education," *The Review of Economic Studies*, 2019, *86* (4), 1487–1526. (cited on p. 5).

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 11 2014, *82* (6), 2295–2326. (cited on pp. 11, 15, and 20).

Cantoni, Davide and Noam Yuchtman, "Historical natural experiments: bridging economics and economic history," in Alberto Bisin and Giovanni Federico, eds., *The Handbook of Historical Economics*, London, UK: Academic Press, 2021, chapter 8, pp. 213–241. (cited on p. 9).

Cappelen, Alexander W., Benjamin Enke, and Bertil Tungodden, "Universalism: Global Evidence," *American Economic Review*, 1 2025, *115*, 43–76. (cited on p. 6).

Carmon, Naomi, "Neighborhood: General," in Neil J. Smelser and Paul B. Baltes, eds., *International Encyclopedia of the Social and Behavioral Sciences*, Oxford: Pergamon, 2001, pp. 10488–10494. (cited on p. 28).

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Manipulation testing based on density discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261. (cited on p. 12).

Cattaneo, Matias D., Rocío Titiunik, and Ruiqi Yu, "Estimation and Inference in Boundary Discontinuity Designs: Location-Based Methods," *arXiv Working Paper 2505.05670*, 5 2025. (cited on pp. 12, 16, and 17).

_ , Rocío Titiunik, and Ruiqi (Rae) Yu, "Estimation and inference in boundary discontinuity designs: Distance-based methods," *Journal of Econometrics*, 7 2026, *256*, 106266. (cited on p. 11).

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole, and Nils Wernerfelt, "Social capital I: measurement and associations with economic mobility," *Nature*, 8 2022, *608* (7921), 108–121. (cited on pp. 5 and 47).

_ , Nathaniel Hendren, and Lawrence F. Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 4 2016, *106* (4), 855–902. (cited on p. 5).

_ , Rebecca Diamond, Thomas B. Foster, Lawrence F. Katz, Sonya R. Porter, Matthew Staiger, and Laura Tach, "Creating High-Opportunity Neighborhoods: Evidence from the HOPE VI Program," Working Paper, Opportunity Insights January 2026. (cited on pp. 5 and 48).

_ , Will Dobbie, Benjamin Goldman, Sonya R Porter, and Crystal S Yang, "Changing Opportunity: Sociological Mechanisms Underlying Growing Class Gaps and Shrinking Race Gaps in Economic Mobility," *The Quarterly Journal of Economics*, 1 2026. (cited on p. 1).

Chyn, Eric, "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children," *American Economic Review*, 10 2018, *108* (10), 3028–56. (cited on pp. 5 and 43).

_ **and Lawrence F. Katz**, "Neighborhoods Matter: Assessing the Evidence for Place Effects," *Journal of Economic Perspectives*, 11 2021, *35* (4), 197–222. (cited on pp. 1, 5, and 13).

_ **, Kareem Haggag, and Bryan A. Stuart**, "The Effects of Racial Segregation on Intergenerational Mobility: Evidence from Historical Railroad Placement," *American Economic Journal: Applied Economics*, 4 2026, *18*, 34–60. (cited on pp. 10 and 45).

Cohn, Alain and Michel André Maréchal, "Priming in Economics," *Current Opinion in Psychology*, 2016, *12*, 17–21. (cited on p. 34).

Davis, Donald R. and David E. Weinstein, "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 11 2002, *92* (5), 1269–1289. (cited on pp. 5 and 9).

de Arregui, Domingo Lázaro, *Descripción de la Nueva Galicia* 1946. (cited on pp. 8 and 10).

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78* (6), 1863–1903. (cited on pp. 5, 10, and 11).

Dickinson, Robert E., *The West European City: A Geographical Interpretation*, London: Routledge & Paul, 1961. (cited on p. 28).

Domínguez, Magdalena, Hans Grönqvist, and Torsten Santavirta, "Effects of Neighborhood Labeling on Student Performance and Sorting," Discussion Paper 38, Helsinki Graduate School of Economics 2025. (cited on p. 48).

Durante, Ruben, Nicola Mastrorocco, Luigi Minale, and James M. Snyder, "Unpacking Social Capital," *The Economic Journal*, 3 2025, *135* (667), 773–807. (cited on pp. 6 and 26).

Durlauf, Steven N., "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1996, *1* (1), 75–93. (cited on p. 1).

_ , "Neighborhood Effects," *Handbook of Regional and Urban Economics*, 1 2004, *4*, 2173–2242. (cited on p. 5).

Enke, Benjamin, "Kinship, cooperation, and the evolution of moral systems," *Quarterly Journal of Economics*, 2019, *134*, 953–1019. (cited on p. 6).

_ **, Ricardo Rodríguez-Padilla, and Florian Zimmermann**, "Moral Universalism and the Structure of Ideology," *The Review of Economic Studies*, 7 2023, *90*, 1934–1962. (cited on p. 37).

Fryer, Roland G. and Paul Torelli, "An empirical analysis of 'acting white'," *Journal of Public Economics*, 6 2010, *94* (5-6), 380–396. (cited on p. 5).

_ **and Steven D. Levitt**, "The Causes and Consequences of Distinctively Black Names," *The Quarterly Journal of Economics*, 2004, *119* (3), 767–805. (cited on pp. 24 and 25).

Goldewijk, Kees Klein, Arthur Beusen, Jonathan Doelman, and Elke Stehfest, "Anthropogenic land use estimates for the Holocene - HYDE 3.2," *Earth System Science Data*, 12 2017, *9* (2), 927–953. (cited on p. 9).

Goldin, Claudia and Lawrence F Katz, "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910-1940," *The Journal of Interdisciplinary History*, 1999, *29* (4), 683–723. (cited on p. 6).

Govind, Yajna, Jack Melbourne, Sara Signorelli, and Edith Zink, "The Making of a Ghetto: Place-Based Policies, Labeling, and Impacts on Neighborhoods and Individuals," Discussion Paper 83/25, ROCKWOOL Foundation Berlin 2025. (cited on p. 48).

Graham, Richard, *Independence in Latin America. Contrast and comparisons*, University of Texas Press, 2013. (cited on p. 7).

Grosfeld, Irena and Ekaterina Zhuravskaya, "Cultural vs. economic legacies of empires: Evidence from the partition of Poland," *Journal of Comparative Economics*, 2 2015, *43*, 55–75. (cited on p. 5).

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The Role of Social Capital in Financial Development," *American Economic Review*, 6 2004, *94* (3), 526–556. (cited on p. 6).

Gutiérrez, Felipe Castro, "El origen y conformación de los barrios de indios," in Felipe Castro Gutiérrez, ed., *Los indios y las ciudades de Nueva España*, 1ª ed., Mexico: UNAM, Instituto de Investigaciones Históricas, 2010, pp. 105–122. (cited on pp. 7 and 9).

Haaland, Ingar, Christopher Roth, Stefanie Stantcheva, and Johannes Wohlfart, "Understanding Economic Behavior Using Open-Ended Survey Data," *Journal of Economic Literature*, 12 2025, *63* (4), 1244–1280. (cited on p. 34).

Hanlon, W.Walker and Stephan Heblich, "History and urban economics," *Regional Science and Urban Economics*, 5 2022, *94*, 103751. (cited on pp. 5 and 19).

Heblich, Stephan, Alex Trew, and Yanos Zylberberg, "East-Side Story: Historical Pollution and Persistent Neighborhood Sorting," *Journal of Political Economy*, 5 2021, *129* (5), 1508–1552. (cited on pp. 5 and 10).

Hornbeck, Richard, Guy Michaels, and Ferdinand Rauch, "Identifying Agglomeration Shadows: Long-run Evidence from Ancient Ports," Technical Report 2025. (cited on p. 10).

IMEPLAN, "Plan de Ordenamiento Territorial Metropolitano del Area Metropolitana de Guadalajara," Technical Report, Instituto Metropolitano de Planeación, Guadalajara 2016. (cited on p. 8).

Jack, B. Kelsey, Kathryn McDermott, and Anja Sautmann, "Multiple price lists for willingness to pay elicitation," *Journal of Development Economics*, 11 2022, *159*, 102977. (cited on pp. 31, 32, 33, and 35).

Jackson, Matthew O., "Inequality's Economic and Social Roots: The Role of Social Networks and Homophily," in Victor Chernozhukov, Johannes Hörner, Eliana La Ferrara, and Iván Werning, eds., *Advances in Economics and Econometrics: Twelfth World Congress*, Cambridge University Press, 2026, pp. 215–258. (cited on pp. 1 and 5).

Jaramillo-Molina, Máximo Ernesto, ""De la Calzada para allá": desigualdad, segregación y estigmatización en el Área Metropolitana de Guadalajara, México," *Lavboratorio*, 2022, *32* (2), 11–55. (cited on p. 10).

Jaramillo-Molina, Máximo and Alejandra Saucedo, ""De la Calzada para allá": Fronteras materiales y simbólicas de desigualdad, segregación y estigmatización en la ciudad de Guadalajara," in "Pobreza y exclusión: diagnósticos y estraegias" 2016, pp. 77–106. (cited on p. 10).

Kamel, Donia, "Between Arab and White: Syrians and the U.S. Naturalization Law," Working Paper, Paris School of Economics 2026. First version November 2024. (cited on p. 24).

Katz, Lawrence F., "Presidential Address: Neighborhood Effects and Missing Markets for Opportunity," *American Economic Review*, 2026, *116* (9), 3197–3222. (cited on pp. 1 and 5).

Keele, Luke and Rocío Titiunik, "Geographic Boundaries as Regression Discontinuities," *Political Analysis*, 1 2015, *23* (1), 127–155. (cited on p. 12).

_ , Scott Lorch, Molly Passarella, Dylan Small, and Rocío Titiunik, "An Overview of Geographically Discontinuous Treatment Assignments with an Application to Children's Health Insurance," in Matias D. Cattaneo and Juan Carlos Escanciano, eds., *Regression Discontinuity Designs: Theory and Applications*, Vol. 38 of *Advances in Econometrics*, Emerald Publishing Limited, 2017, pp. 147–194. (cited on p. 12).

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 1 2007, *75* (1), 83–119. (cited on p. 5).

Knack, Stephen and Philip Keefer, "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics*, 11 1997, *112* (4), 1251–1288. (cited on p. 6).

Knight, Alan, *Mexico. The colonial era*, Cambridge University Press, 2002. (cited on p. 7).

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355. (cited on p. 12).

Lee, Sanghoon and Jeffrey Lin, "Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income," *The Review of Economic Studies*, 1 2018, *85* (1), 663–694. (cited on p. 9).

Lin, Jeffrey and Ferdinand Rauch, "What future for history dependence in spatial economics?," *Regional Science and Urban Economics*, 5 2022, *94*, 103628. (cited on pp. 5 and 19).

Lowes, Sara, "Culture in Historical Political Economy," in "The Oxford Handbook of Historical Political Economy," Oxford University Press, 1 2023, pp. 887–924. (cited on p. 5).

_ , "Kinship Structure and the Family: Evidence from the Matrilineal Belt," NBER Working Paper 30509, National Bureau of Economic Research 2023. (cited on p. 6).

_ **and Eduardo Montero**, "The Legacy of Colonial Medicine in Central Africa," *American Economic Review*, 4 2021, *111* (4), 1284–1314. (cited on p. 6).

_ **, Nathan Nunn, James A. Robinson, and Jonathan L. Weigel**, "The Evolution of Culture and Institutions: Evidence From the Kuba Kingdom," *Econometrica*, 7 2017, *85* (4), 1065–1091. (cited on p. 10).

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu, "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity," *American Economic Review*, 5 2013, *103* (3), 226–31. (cited on p. 5).

Maloney, William F. and Felipe Valencia-Caicedo, "The Persistence of (Subnational) Fortune," *The Economic Journal*, 12 2016, *126* (598), 2363–2401. (cited on p. 9).

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 7 1993, *60* (3), 531–542. (cited on pp. 13 and 30).

Matías, Alejandro Solís, *Analco*, Gobierno del Estado de Jalisco, 1986. (cited on p. 9).

_ , "San José de Analco en 1907," *Cuadernos de Estudios Jaliscienses*, 1992. (cited on pp. 9 and 10).

McCartan, Cory, Jacob R. Brown, and Kosuke Imai, "Measuring and Modeling Neighborhoods," *American Political Science Review*, 11 2024, *118* (4), 1966–1985. (cited on pp. 5 and 31).

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2 2008, *142* (2), 698–714. (cited on p. 12).

Michalopoulos, Stelios and Melanie Meng Xue, "Folklore," *The Quarterly Journal of Economics*, 2021, *136* (4), 1993–2046. (cited on p. 5).

_ , **Elie Murard, Elias Papaioannou, and Seyhun Orcan Sakalli**, "Uprootedness, Human Capital, and Skill Transferability," NBER Working Paper 33586, National Bureau of Economic Research March 2025. (cited on p. 5).

Müller, Ulrich K. and Mark W. Watson, "Spatial Unit Roots and Spurious Regression," *Econometrica*, 9 2024, *92* (5), 1661–1695. (cited on p. 22).

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson, "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," *The Review of Economic Studies*, 5 2022, *89* (3), 1557–1592. (cited on pp. 5 and 43).

Nunn, Nathan, "Culture and the Historical Process," *Economic History of Developing Regions*, 6 2012, *27* (sup1), S108–S126. (cited on p. 5).

_ , "Historical Development," in "Handbook of Economic Growth" 2014, pp. 347–402. (cited on p. 5).

_ , "The historical roots of economic development," *Science*, 3 2020, *367* (6485). (cited on p. 5).

_ **and Leonard Wantchekon**, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 12 2011, *101* (7), 3221–52. (cited on p. 5).

Oh, Suanna, "Does Identity Affect Labor Supply?," *American Economic Review*, 2023, *113* (8), 2055–2083. (cited on p. 6).

Portes, Alejandro, "Social Capital: Its Origins and Applications in Modern Sociology," *Annual Review of Sociology*, 8 1998, *24*, 1–24. (cited on p. 6).

Ramos-Toro, Diego, "Social Exclusion and Social Preferences: Evidence from Colombia's Leper Colony," *American Economic Review*, 5 2023, *113* (5), 1294–1333. (cited on pp. 6 and 9).

Reguillo, Rossana, "Semantizarás el territorio: los vecinos de Analco y las explosiones de 1992 en Guadalajara," *Alteridades*, 1998, *8* (15), 35–44. (cited on pp. 10 and 41).

Rojas-Ampuero, Fernanda and Felipe Carrera, "Sent Away: Displacement, Neighborhoods, and Children's Outcomes under Slum Clearance Policies," 6 2026. Working paper. (cited on p. 5).

Rossignol, Etienne Le and Sara Lowes, "Mobility, Competition, and the Scope of Trust: Evidence from Transhumant Pastoralism," May 2026. Working paper. (cited on p. 6).

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley, "Assessing "Neighborhood Effects": Social Processes and New Directions in Research," *Annual Review of Sociology*, 2002, *28*, 443–478. (cited on pp. 1 and 35).

Sen, Amartya, "Social Exclusion: Concept, Application, and Scrutiny," Social Development Papers 1, Asian Development Bank, Manila, Philippines 6 2000. (cited on p. 9).

Shayo, Moses, "Social Identity and Economic Policy," *Annual Review of Economics*, 8 2020, *12* (1), 355–389. (cited on p. 6).

Smith, Michael E., "The archaeological study of neighborhoods and districts in ancient cities," *Journal of Anthropological Archaeology*, 6 2010, *29* (2), 137–154. (cited on p. 35).

Tabellini, Guido, "The Scope of Cooperation: Values and Incentives," *Quarterly Journal of Economics*, 8 2008, *123*, 905–950. (cited on p. 6).

Tanck de Estrada, Dorothy, *Pueblos de indios y educación en el México colonial, 1750–1821*, Mexico City: El Colegio de México, Centro de Estudios Históricos, 1999. (cited on p. 10).

_ , *Atlas ilustrado de los pueblos de indios de la Nueva España, 1800*, Fomento Cultural Banamex A C, 2005. (cited on p. 8).

Topa, Giorgio and Yves Zenou, "Neighborhood and Network Effects," in "Handbook of Regional and Urban Economics" 2015, pp. 561–624. (cited on p. 5).

Valencia-Caicedo, Felipe, "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America," *The Quarterly Journal of Economics*, 2 2019, *134* (1), 507–556. (cited on p. 5).

Valenzuela-Casasempere, Pablo, "Displacement and Infrastructure Provision: Evidence from the Interstate Highway System," Technical Report 2025. (cited on p. 10).

Voth, Hans-Joachim, "Persistence – myth and mystery," *The Handbook of Historical Economics*, 1 2021, pp. 243–267. (cited on p. 5).

Weigand, Matthias, Cathrin Mohr, and Davide Cantoni, "Identity and Institutional Change: Evidence from First Names in Germany, 1700–1850," 8 2025. (cited on p. 24).

Woo-Mora, L. Guillermo, "Unveiling the Cosmic Race: Skin Tone and Intergenerational Economic Disparities in Latin America and the Caribbean," *Journal of Development Economics*, 2026, *179*, 103594. (cited on p. 73).

_ **and Pedro Torres-López**, "New Echoes of Deep Mexico? The Persistent Elements of Indigenous Identities," 2026. (cited on p. 13).

A Historical Natural Experiment

This appendix collects the supporting evidence for Section 2, in the order the section draws on it. Appendix A.1 locates Guadalajara in the national distribution of income, schooling, and school attendance, and Appendix A.2 traces attendance back to 1970. Appendix A.3 documents the absence of pre-colonial agglomerations and locational fundamentals that could have driven settlement, and Appendix A.4 traces the persistence of the urban form from the colonial creek to present-day Calzada Independencia. Appendix A.5 reports the realized university wage premium. Appendix A.6 reports the 1821 socio-demographic comparison across banks, Appendix A.7 the skin-tone evidence, and Appendix A.8 the 1882 cadaster and the early capitalization of exclusion into land values. Appendix A.9 documents where the institutions named on the city's general maps stood relative to the creek between 1800 and 1908, and the population living on each bank.

A.1 Guadalajara in the national distribution

Figure A.1 plots, for each claim in the context paragraph of Section 2, the distribution of the outcome across Mexico's urban municipalities, with Guadalajara and the national value marked. The municipality sits in the 98th percentile of household income and the 95th of mean schooling, and its school attendance exceeds the national rate at every level beyond primary.

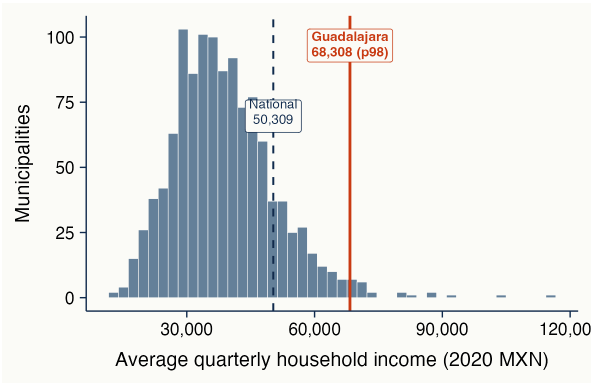
A.2 School attendance over time and where it varies

Figure A.2 extends the attendance comparison of Section 2 back to 1970 using census microdata. Guadalajara exceeds the national rate at ages 12 and above in every census, and the advantage is concentrated at the post-primary margins. At ages 6–11 the municipality led the national rate by six points in 1970, but the gap closed as attendance approached universal coverage nationwide, and in 2010 and 2020 Guadalajara sits marginally below it (93.7 against 94.5 percent in 2020). At ages 18–24 its lead instead widened, from 21.5 against 12.0 percent in 1970 to 42.7 against 32.4 percent in 2020.

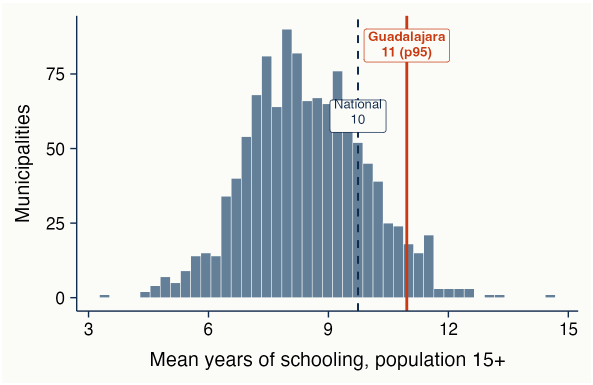
Figure A.3 asks where that variation lives. Writing y for whether a young person aged 12–17 attends school, the variance splits exactly into three nested parts: differences between municipalities, differences between households of the same municipality, and differences between siblings of the same household. The municipality accounts for very little and for steadily less: 8.8 percent of the variance in 1990 and 3.7 percent in 2020. Almost all of the remainder sits between households living in the same municipality, a component that rose from 60.5 to 78.7 percent over the same period. Educational inequality in Mexico thus moved away from differences across places and toward differences across households within a place.

Two further results follow. Observable household characteristics account for only part of the household component. Age, sex, the schooling of the household head or spouse, birth order

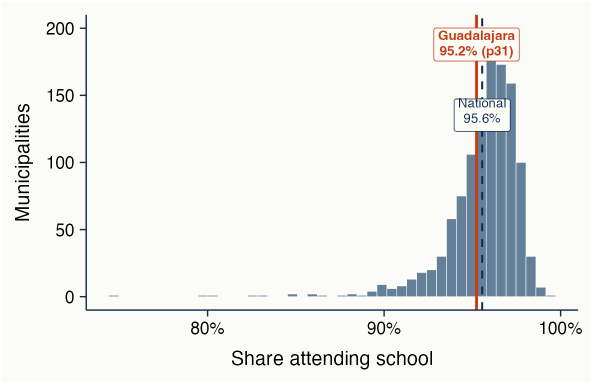
Figure A.1: Guadalajara in the national distribution of income, schooling, and attendance (2020)



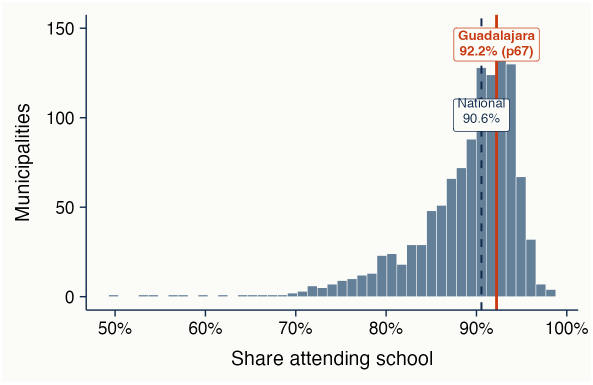
(a) Quarterly household income



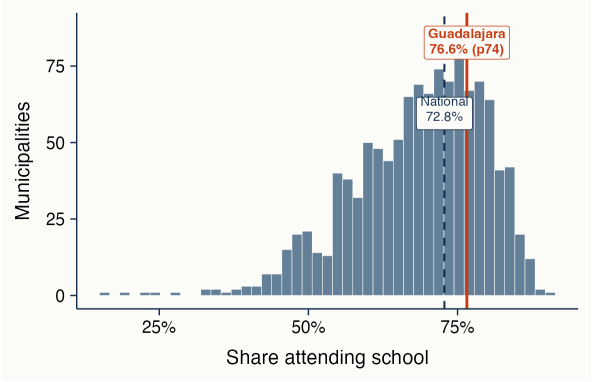
(b) Years of schooling, ages 15+



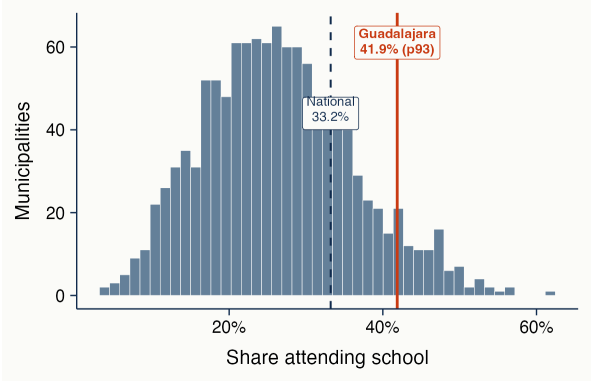
(c) School attendance, ages 6–11



(d) School attendance, ages 12–14



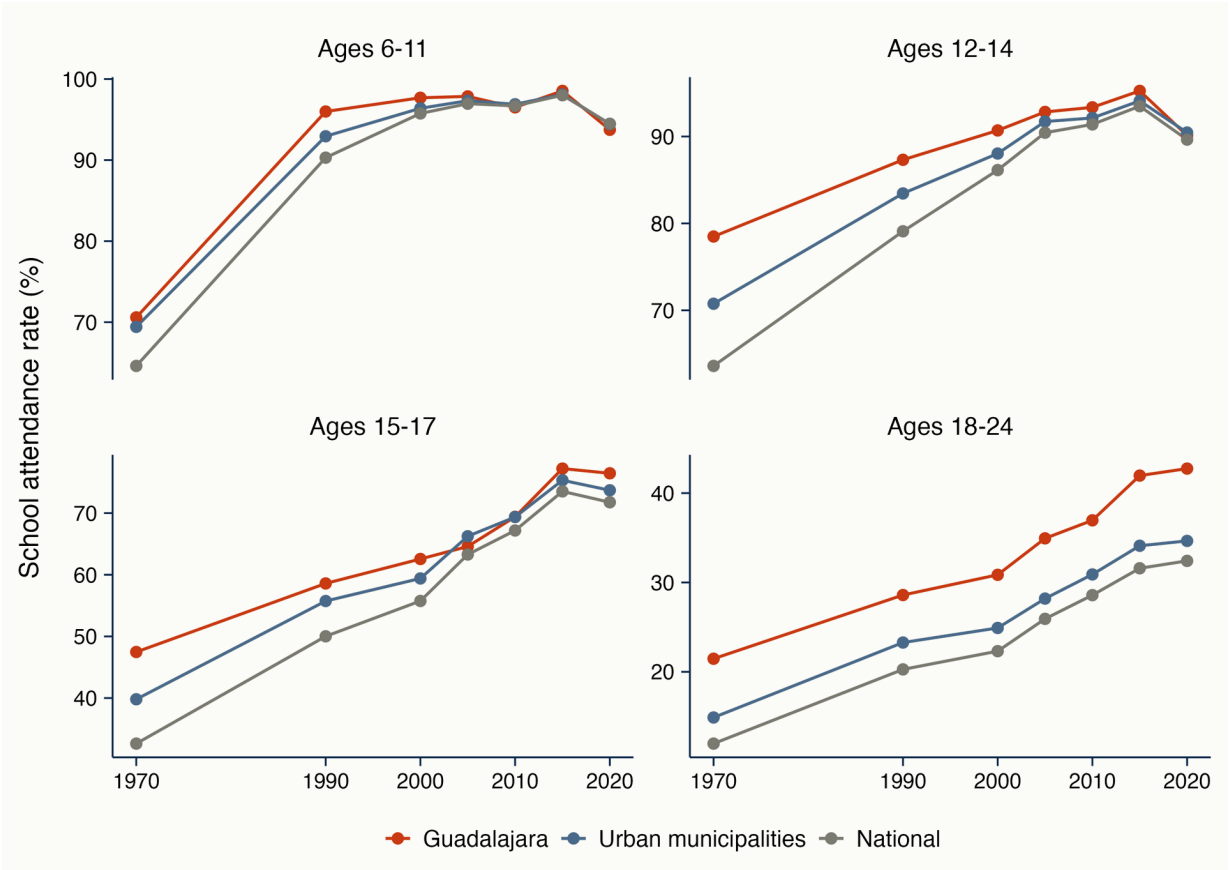
(e) School attendance, ages 15–17



(f) School attendance, ages 18–24

Notes: Each panel plots the distribution of the outcome across the 1,169 urban municipalities of Mexico (municipalities with at least 15,000 residents in the 2020 census). The solid line marks Guadalajara, labeled with its value and its percentile within the urban distribution; the dashed line marks the national value. Household income is the 2020 average quarterly household income in pesos at the municipal level; schooling and attendance come from the municipal statistics of the 2020 *Censo de Población y Vivienda* (ITER). Attendance at ages 6–11 is near-universal everywhere, so Guadalajara's advantage appears only at the post-primary levels, as stated in the text.

Figure A.2: School attendance by age group, 1970–2020



Notes: Share of individuals in each age group reported as attending school, computed from the IPUMS International samples of the Mexican censuses of 1970, 1990, 2000, 2010 and 2020 and the intercensal counts of 2005 and 2015, weighted by person weights. "Urban municipalities" are the harmonized municipalities whose population in urban localities reaches at least 50 percent in the corresponding year, excluding Guadalajara; "National" pools all municipalities. The 1995 count is omitted because its sampling design and geographic coverage differ from those of the censuses. The 1960 sample does not record school attendance.

and sibship size jointly explain 13.7 percent of the total variance in 2020, and net of them the municipality share falls to 1.6 percent while the household share rises to 80.2 percent. The modest explanatory share should not be read as parental schooling being unimportant, since it shifts attendance substantially: 57.5 percent of 12–17 year-olds attend school where the head or spouse has at most five years of schooling, against 92.4 percent where that figure reaches twelve years. Parental schooling alone accounts for 8.3 percent of the variance because attendance is binary and most of its dispersion is idiosyncratic. And the household component is not merely the sum of individual differences. Among households with at least two children in the age range, where the split is separately identified, 59.0 percent of the variance lies between households and 36.2 percent between siblings within a household.

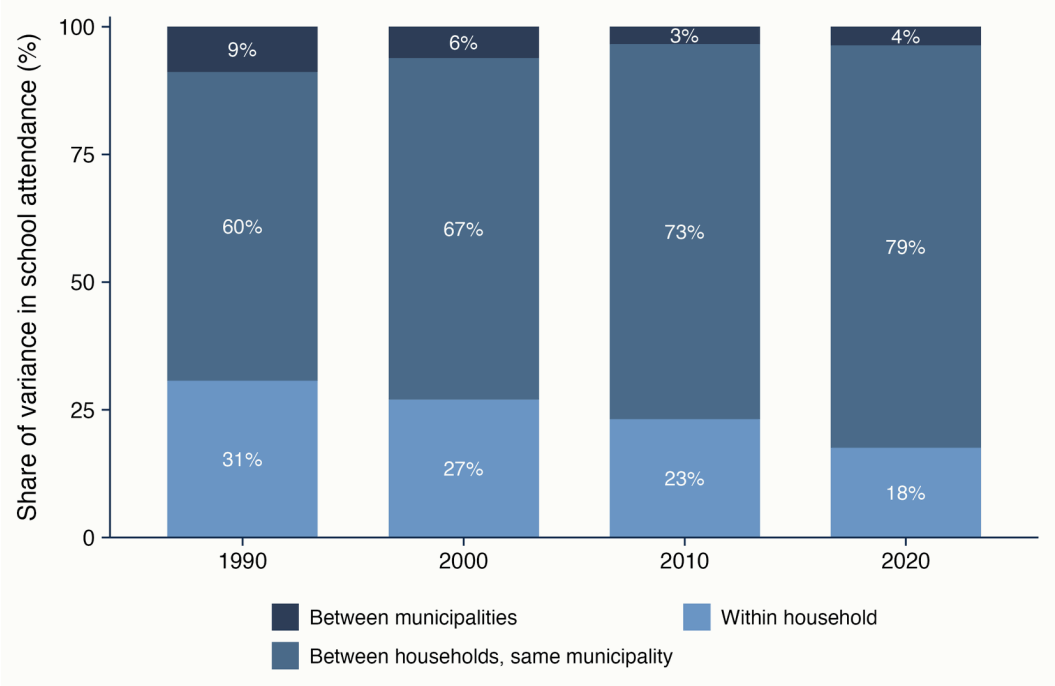
The exercise cannot go further than the municipality, because the census records no geography below it. The between-household component therefore pools whatever distinguishes households from whatever distinguishes the neighborhoods they live in. Separating the two requires variation in neighborhood below the municipality. The boundary in Section 3 supplies it, comparing households a few blocks apart within a single municipality, on either side of a line that other determinants of schooling vary smoothly across. That design compares different households rather than following one household across neighborhoods, so it does not by itself separate place effects from the composition of residents, a limitation Section 3 discusses.

A.3 Pre-colonial population and locational fundamentals

The natural-experiment argument requires that nothing already distinguished the two banks of the SJdD creek when the settlements were founded in 1542 — that neither a pre-colonial population center nor an exogenous geographic feature dictated which bank each community occupied. This subsection presents the evidence behind that claim, first at the scale of the valley and then at the scale of the boundary itself.

Pre-colonial population. Figure A.4 shows pre-colonial population density across Mesoamerica circa 1500, drawing on the historical reconstruction of Goldewijk et al. (2017). The Anahuac Valley (Panel b) — the heartland of the Mexica state and home to Tenochtitlán — concentrated densely settled urban polities. The Atemajac Valley (Panel c), where Guadalajara would later be founded, contained dispersed Indigenous populations but no major agglomerations. The valley therefore held no established urban center whose location could have dictated the 1542–1543 allocation of colonizers, allies, and conquered populations across the SJdD creek. That statement concerns the valley rather than the creek. The reconstruction is gridded at five arc-minutes, about nine kilometers at this latitude, so a single cell covers the entire estimation window and cannot speak to settlement or land use at the boundary itself. It also rests on the pre-colonial reconstruction alone. Figure A.5 reports population estimates from Buringh (2018) for Mexico City and Guadalajara from the 16th to 18th centuries, showing that Guadalajara grew slowly

Figure A.3: Where the variance in school attendance lies, 1990–2020



Notes: Nested decomposition of the variance of an indicator for attending school among individuals aged 12–17 who are children of the household head, using the IPUMS International samples of the 1990, 2000, 2010 and 2020 Mexican censuses and weighting by person weights. The three components are $\mathrm{Var}_m(\bar{y}_m)$, $\mathbb{E}_m[\mathrm{Var}_h(\bar{y}_h)]$ and $\mathbb{E}_h[\mathrm{Var}(y)]$, where m indexes harmonized municipalities and h households nested within them; they sum to the total variance by construction. Households with a single child in the age range contribute no within-household variance, so their individual variation enters the household component; restricting to households with two or more such children, which identifies the split separately, yields 4.8, 59.0 and 36.2 percent in 2020. The census provides no geography below the municipality, so the household component contains both household and neighborhood differences.

relative to Mexico City and from a low initial base, but that series begins after the foundation and documents how the asymmetry persisted rather than what preceded it. What the design requires is not an empty landscape but the absence of a discontinuity at the creek, which Table A.1 examines directly.

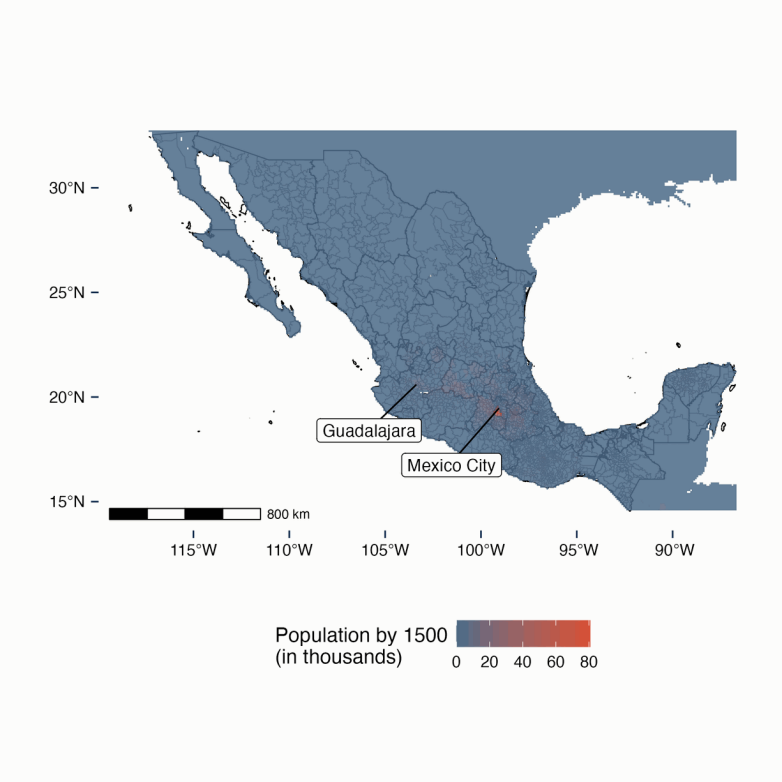
Locational fundamentals. A complementary concern is that, even absent pre-colonial agglomerations, geographic features could have determined which side of the SJdD creek was settled by colonizers and which by the conquered. Figure A.6 maps altitude at the 250 × 250 m grid-cell level across the foundational sites, using INEGI's high-resolution raster data. The terrain is essentially uniform: there is no elevation differential between Spaniard Guadalajara and Mexicaltzingo (west bank) and Analco (east bank), and no topographic feature that would have imposed a particular settlement pattern. Table A.1 formalizes this assessment. I implement a regression discontinuity design at the grid-cell level within a 1.5-km buffer of the SJdD creek, with the cell's signed orthogonal distance to the creek as the running variable, and test for differences across banks in six locational fundamentals: altitude, slope, mean temperature, precipitation, humidity, and thermal variation. None of the six estimates is statistically significant, and every sharpened FDR-adjusted q-value (Anderson, 2008) remains above 0.6 after correction for multiple testing. The two topographic tests are the informative ones, and they are precise. Altitude and slope come from INEGI's continuous elevation model at 15 m native resolution, a direct measurement that resolves the boundary, and their confidence intervals exclude differences larger than 0.03 and 0.13 standard deviations. The four climate layers are interpolated long-run normals for 1902–2011, so smoothness across a three-kilometer window is in part a property of the interpolation rather than an independent finding, and two of them leave intervals reaching about 0.45 standard deviations. No layer measures hydrology, and the one hydrological feature at the boundary is the creek itself, which defines the assignment and therefore cannot serve as a balance test. These estimates bound differences across banks rather than establish equivalence. What they show is no topographic feature large enough to have dictated which bank each community occupied, which is consistent with the documentary record that the allocation was administrative.

A.4 Persistence of the urban form: from creek to Calzada

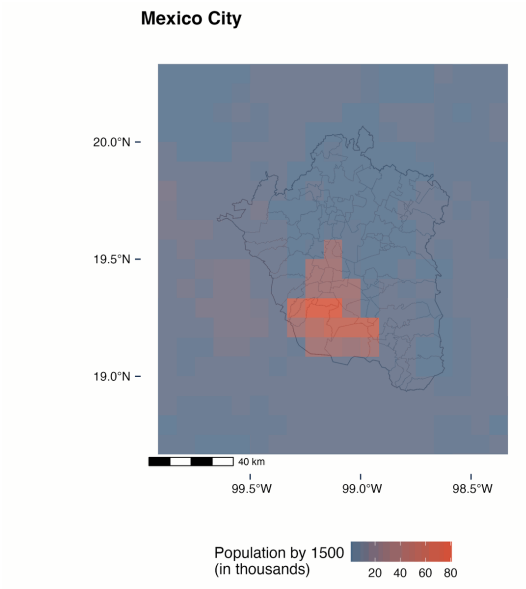
Unlike many Mexican cities, where 19th- and 20th-century reorganizations erased earlier divisions, Guadalajara kept the initial 1542 structure as the basis of its long-run urban form. This subsection documents that persistence visually.

Limited urban expansion before the 20th century. The 1621 account of priest Domingo Lázaro de Arregui describes the SJdD creek as a modest watercourse crossed by multiple bridges (Lázaro de Arregui, 1946). Figure A.7, panel (a) reproduces a 1908 map: nearly four centuries after the foundation, the city's footprint remained largely contained within the colonial core,

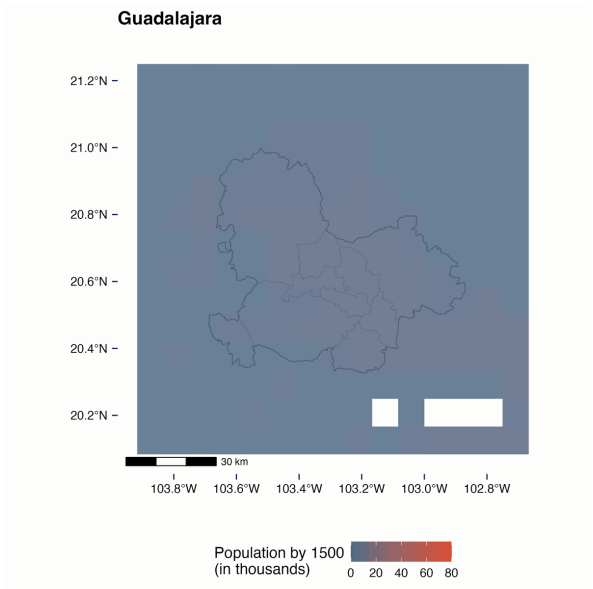
Figure A.4: Pre-colonial population density estimates in Mesoamerica



(a) Pre-colonial population density across Mexico



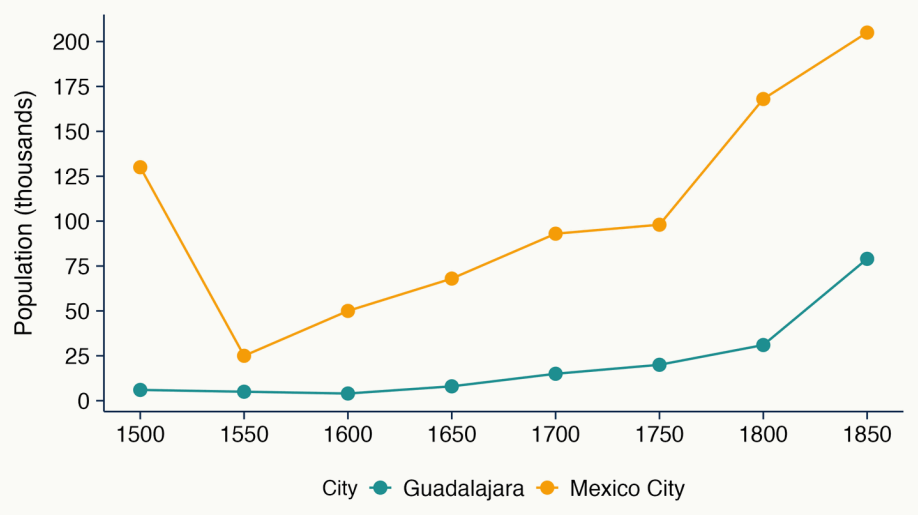
(b) Anahuac Valley (Central Mexico)



(c) Atemajac Valley (Guadalajara region)

Notes: Panel (a) shows pre-colonial population density across present-day Mexico circa 1500 CE, based on the historical land-use reconstruction of Goldewijk et al. (2017). Panels (b) and (c) zoom into the Anahuac Valley (the seat of Tenochtitlán and the Mexica state) and the Atemajac Valley (the future site of Guadalajara), respectively. The Anahuac Valley contained densely settled urban polities at the time of Spanish contact, while the Atemajac Valley hosted dispersed Indigenous populations but no major agglomerations within the area where Spaniard Guadalajara, Mexicaltzingo, and Analco would later be founded. Local Indigenous groups including the Cazcanes, Cocas, and Tecuexes inhabited the broader region, but no urban center existed at the scale relevant for the 1542 foundation.

Figure A.5: Colonial population estimates for Mexico City and Guadalajara



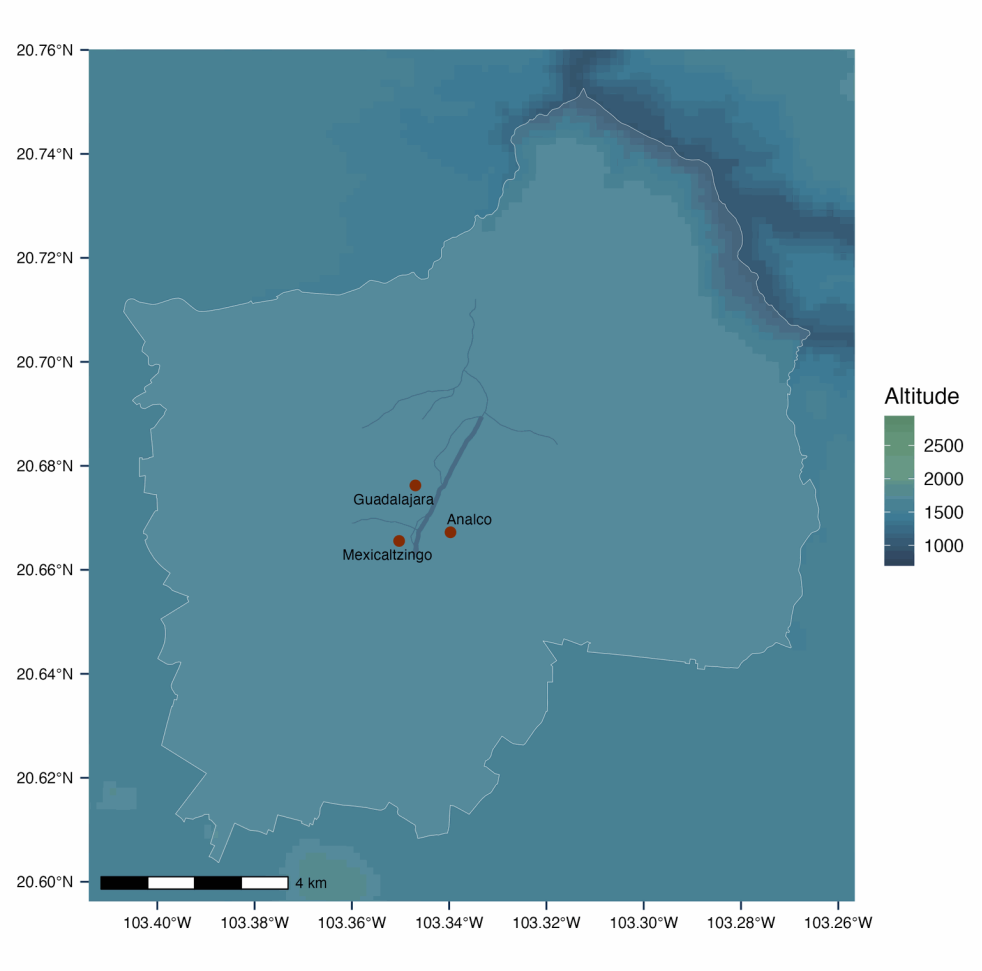
Notes: Population estimates for Mexico City and Guadalajara from the 16th to 18th centuries, based on the historical urban-population reconstructions of Buringh (2018). The persistent population gap between the two cities — Mexico City consistently 5–10 times larger than Guadalajara through the colonial period — reflects the absence of pre-colonial urban agglomeration in the Atemajac Valley and the slower demographic growth of New Galicia relative to central New Spain. The figure rules out the possibility that the 1542 foundation of Guadalajara built upon a substantial pre-existing urban base.

Table A.1: Balance in locational fundamentals

	RDD estimate	SE	95% CI	p-value	FDR q-value
Altitude	0.002	0.015	[−0.027, 0.031]	0.910	0.922
Slope	-0.053	0.039	[−0.128, 0.023]	0.174	0.646
Temperature	-0.004	0.038	[−0.078, 0.070]	0.922	0.922
Precipitation	-0.163	0.144	[−0.446, 0.120]	0.258	0.646
Humidity	-0.036	0.055	[−0.143, 0.071]	0.510	0.766
Thermal variation	0.154	0.156	[−0.152, 0.460]	0.323	0.646

Notes: Estimates from a spatial regression discontinuity design at the grid-cell level, using INEGI's 250 × 250 m raster data on locational fundamentals within a 1.5-km buffer of the San Juan de Dios creek (today Calzada Independencia). The running variable is the cell's signed orthogonal distance to the historical boundary, with positive values on the eastern (Analco) side. Each row reports the estimated discontinuity for a different fundamental: altitude (m), slope (degrees), mean annual temperature (°C), annual precipitation (mm), humidity (Lang aridity index, precipitation over temperature), and thermal variation (°C). All variables are standardized within the analysis window. Standard errors and 95% confidence intervals are robust and bias-corrected, and bandwidth selection and inference follow Calonico et al. (2014). Sharpened FDR-adjusted q-values (Anderson, 2008) are computed across the six fundamentals. None of the estimates is statistically significant, and all q-values remain above 0.6 after correction. The intervals bound rather than exclude differences across banks. They are tightest for altitude and slope, which derive from a 15 m elevation model, and widest for precipitation and thermal variation, which come from interpolated climate normals.

Figure A.6: Altitude across the historical boundary: 250 × 250 m grid-cell raster



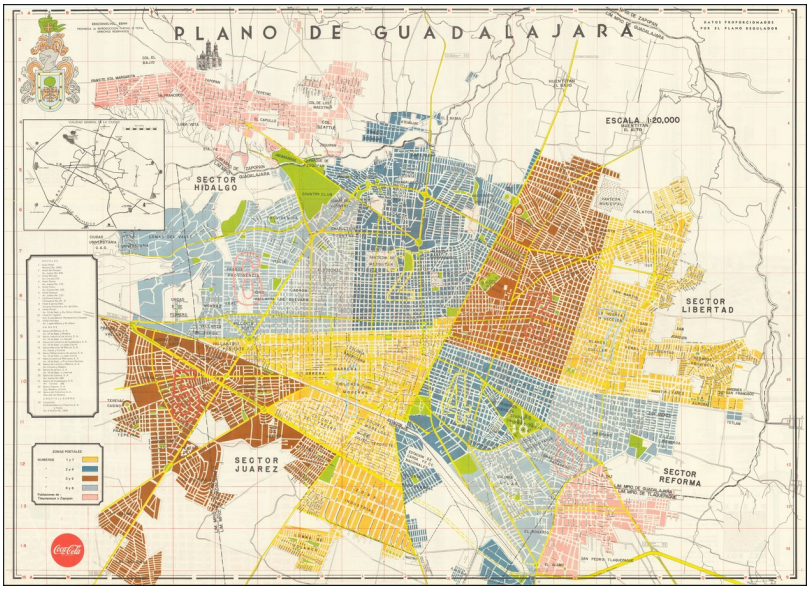
Notes: Altitude in meters above sea level at the 250 × 250 m grid-cell level, from INEGI's high-resolution digital elevation model. The figure overlays the historical boundary (San Juan de Dios creek, today Calzada Independencia) and the approximate locations of the three foundational settlements (Spaniard Guadalajara, Mexicaltzingo, and Analco). Color shading indicates altitude, with no systematic differential across the SJdD creek. The visual impression is confirmed formally in Table A.1.

and the creek remained a legible east–west divide. Panel (b) reproduces a 1972 map: by the second half of the 20th century the metropolitan area had expanded dramatically following mid-century industrialization (Stephen H. Haber, 1987), but the colonial core retained its centrality and Calzada Independencia continued to demarcate the east–west structure.

Figure A.7: Urban expansion of Guadalajara across the 20th century



(a) Limited urban expansion by the early 20th century (1908)



(b) Metropolitan expansion by the 1970s

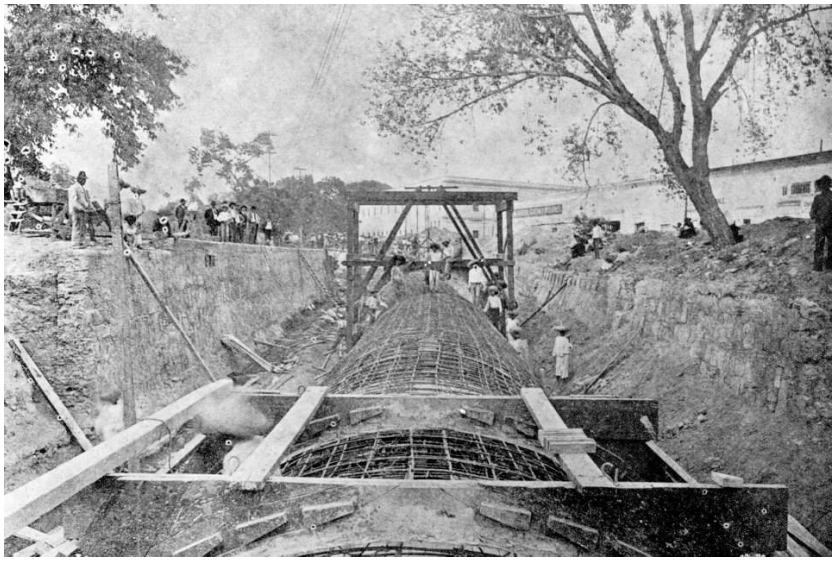
Notes: Panel (a) reproduces a 1908 cartographic representation of Guadalajara: the footprint remained largely contained within the colonial core nearly four centuries after the 1542 foundation, and the SJdD creek divided Spaniard Guadalajara and Mexicaltzingo on the western bank from Analco on the eastern bank. Panel (b) reproduces a 1972 map from the Instituto de Información Estadística y Geográfica (IIEG) at the height of Mexico's mid-century industrialization. The metropolitan footprint expanded by an order of magnitude, but the colonial core retained its centrality, and Calzada Independencia — which by 1972 had replaced the SJdD creek — continued to demarcate the east–west structure. Map sources: IMEPLAN (2016) and the IIEG historical map collection.

From creek to avenue. Figure A.8 documents the transformation. Panel (a) shows the SJdD creek as a still-visible watercourse in the early 1900s; between 1903 and 1911 municipal authorities piped it underground and built a major north–south avenue along its course (panel b, Calzada Independencia). The transformation eliminated the creek as a physical barrier but preserved its position as the city's main east–west axis.

Figure A.8: Transformation of the SJdD creek into Calzada Independencia, 1903–1911



(a) Early 20th-century view of the SJdD creek



(b) Construction of Calzada Independencia along the creek's course

Notes: Panel (a) shows the SJdD creek in the early 20th century, immediately prior to its reorganization; described as a modest watercourse with multiple bridges in 17th-century accounts (Lázaro de Arregui, 1946), it retained its position as the dividing line through the late 19th century. Panel (b) documents the construction of Calzada Independencia along the same north–south axis, with the creek piped underground, between 1903 and 1911 as part of the Porfirian-era modernization of the city. Photographs from municipal historical archives.

The Calzada as a present-day connector, not a barrier. Figure A.9 shows Calzada Independencia in 2020: a surface-level avenue integrated into the city grid, traversed continuously by pedestrians, vehicles, and a bus rapid transit line. It is not the kind of physical barrier studied in the U.S. literature on urban freeways and railroads (Ananat, 2011; Chyn et al., 2026; Bagagli, 2025; Valenzuela-Casasempere, 2025); if anything it functions as the city's main north–south arterial. For identification, this means the spatial RDD exploits a historical line whose physical form has changed repeatedly, rather than a contemporary infrastructural divide imposing its own treatment effect.

A.5 Realized university wage premium, 1990–2020

Figure A.10 estimates the university wage premium over completed high school in Guadalajara and in other urban municipalities and local labor markets, using census microdata from 1990 to 2020. The premium rises until 2000 for men and 2010 for women and declines afterward, ending between 54 and 70 percent in 2020. In almost every year, at both geographic levels, and for both sexes, the realized premium lies below the 100 percent premium that parents in the survey of Section 5 expect at age 30. The exception is men in other urban municipalities in 2000, whose estimated premium reaches 100 percent with a confidence interval extending above it.

A.6 Socio-demographic sorting in 1821

To document the persistence of the colonial division three centuries after the foundation, I rely on the 1821 *padrón* of Guadalajara, conducted in the immediate aftermath of independence and curated and digitized by the Guadalajara Census Project (Anderson and Spike, 2007). The census enumerated 35,287 residents. It numbers its units from 1 to 25 and records entries for 24 of them, with *cuartel* 16 absent, and gives names, age, sex, household membership, ethnicity (*calidad*), marital status, and detailed occupational information. To the best of my knowledge, the 1821 *padrón* of Guadalajara is the most detailed urban dataset available for nineteenth-century Latin America. Figure A.11 compares mean characteristics of residents on either side of the SJdD creek, restricting the sample to the sixteen *cuarteles* immediately adjacent to the boundary — the same spatial window used in the spatial RDD of Section 3. Three patterns stand out. First, residents on the western bank were significantly more likely to hold noble titles (*Don*/*Doña*) and to identify as Spanish. Second, residents on the eastern bank were more likely to identify as Indigenous, to be in interracial marriages, and to hold journeyman or laborer occupations. Third, residential mobility — proxied by the share of migrants — was higher on the eastern bank. The differences survive controls for age and sex, household clustering, and FDR adjustment. The sign of this last pattern is the opposite of the long-term residence gap measured in 2020 (Table B.8). The two measures are two centuries apart and bracket the twentieth-century rural-to-urban migration that reshaped the city, so they describe different populations under different migration regimes rather

Figure A.9: Calzada Independencia as a present-day surface-level avenue



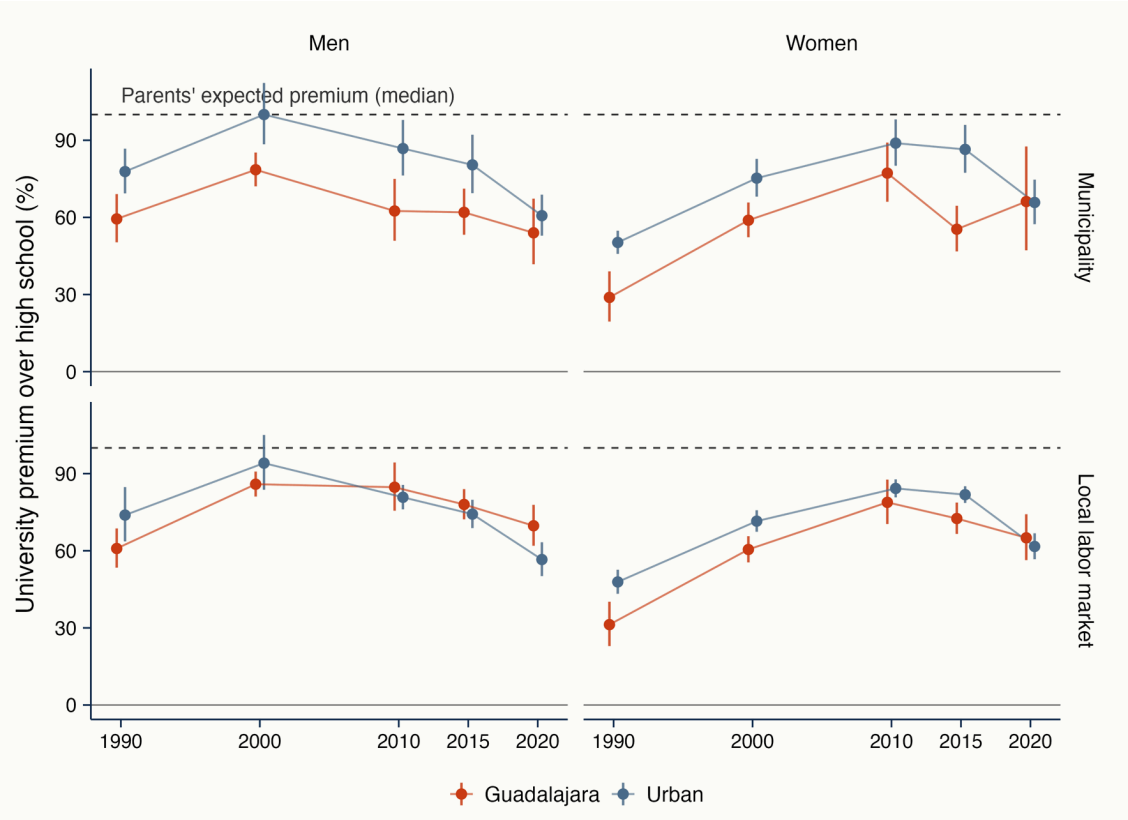
(a) Westward view of Calzada Independencia



(b) Eastward view of Calzada Independencia

Notes: Photographs of Calzada Independencia in 2020. Panel (a) shows the view from the western (Spaniard Guadalajara and Mexicaltzingo) side; panel (b) from the eastern (Analco) side. The avenue is a surface-level urban arterial integrated into the city grid, with frequent pedestrian crossings, continuous commercial activity along both sides, and a bus rapid transit line (*Macrobús*) on its central median. Unlike the urban freeways studied in the U.S. literature (Ananat, 2011; Chyn et al., 2026; Bagagli, 2025; Valenzuela-Casasempere, 2025), Calzada Independencia connects rather than fragments the two banks. The persistence of the east–west divide that the spatial RDD estimates therefore cannot be attributed to the Calzada functioning as a contemporary physical barrier.

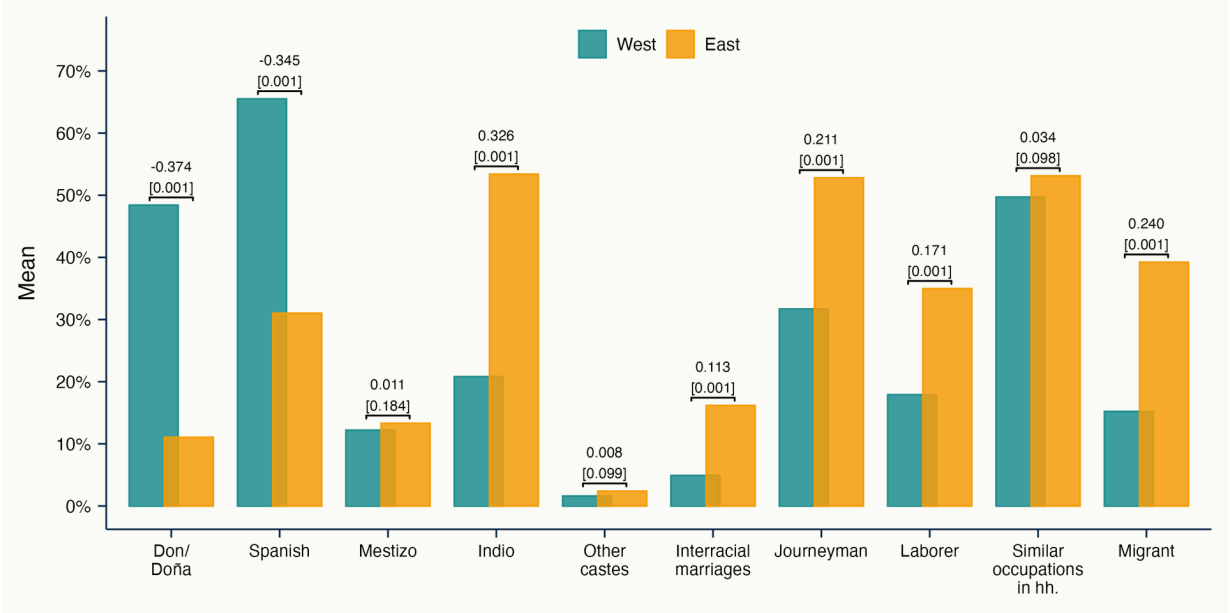
Figure A.10: University wage premium in Guadalajara and urban Mexico, 1990–2020



Notes: Each point reports the university wage premium implied by a Mincer regression of log monthly labor earnings on a university indicator and a quadratic in age, estimated separately by census year, sex, and geographic group on IPUMS International microdata from the 1990, 2000, 2010, and 2020 censuses and the 2015 Intercensal Survey. The sample covers employed workers aged 25–55 whose highest attainment is completed high school or a university degree, restricted to at least 30 weekly hours where hours are recorded; the premium is $100(e^{\beta} - 1)$. The top row compares the municipality of Guadalajara with all other urban municipalities, defined by an urban population share of at least 50 percent in each year; the bottom row repeats the exercise at the level of the local labor markets delimited by Aldeco et al. (2024), comparing Guadalajara's market with other urban markets. Vertical bars show 95 percent confidence intervals, based on heteroskedasticity-robust standard errors for Guadalajara and standard errors clustered by municipality or labor market otherwise. The dashed line marks the median university premium that surveyed parents on both banks expect at age 30 (Section 5). Restricting the 2020 Guadalajara estimate to workers aged 25–35, the age in the survey vignette, yields a realized premium of 57 percent with a 95 percent confidence interval of 44 to 72 percent.

than one series.

Figure A.11: 1821 census: socio-demographic characteristics by bank of the San Juan de Dios creek



Notes: Mean values of selected socio-demographic characteristics in 1821, computed separately for residents of the eight *cuarteles* immediately west of the SJdD creek (Spaniard Guadalajara and Mexicaltzingo) and the eight *cuarteles* immediately east (Analco). Data are from the individual-level 1821 *padrón* compiled by the Guadalajara Census Project (Anderson and Spike, 2007). Variables capture social status (share of household heads titled *Don/Doña*), self-reported ethnicity (share Spanish, share Indigenous), interracial marriage, occupational structure (journeymen and laborers), and residential mobility (share of migrants). Annotations above each pair of bars report the OLS estimate of the east-minus-west difference in means, controlling for age and sex, with standard errors clustered by household in parentheses. Sharpened FDR q-values (Anderson, 2008), adjusted across all variables, are in brackets. The sample is restricted to the sixteen *cuarteles* adjacent to the historical boundary for comparability with the spatial RDD sample. These regressions control for age and sex, not for ethnicity, and the western group includes Mexicaltzingo's Indigenous population, so bank and ethnicity are not the same partition. Table A.2 adds self-reported *calidad* as a control and the non-ethnic gaps remain, so the side of the creek predicts status, occupation, intermarriage and mobility beyond ethnic composition, consistent with the natural experiment isolating colonial-hierarchy standing rather than ethnicity alone.

A.7 Skin tone across the metropolitan area

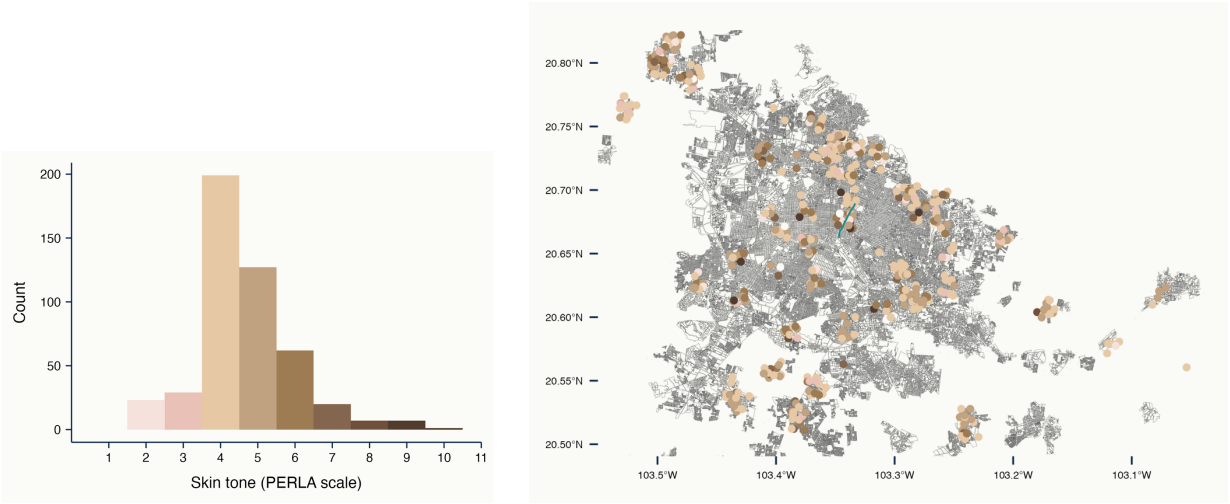
Mexican censuses do not record skin tone, so the ethnic composition of the two banks in Section 3 rests on Indigenous self-identification, household composition and language. Because skin tone rather than Indigenous identity may be the operative distinction (Woo-Mora, 2026), I turn to the 2017 ESRU Survey of Social Mobility (ESRU-EMOVI), which records interviewer-rated skin tone on the PERLA scale together with respondents' coordinates.

Table A.2: 1821 socio-demographic gaps across the boundary, with and without a control for *calidad*

	(1) Full sample	(2) *Calidad* subsample	(3) *Calidad* controlled	Observations
Don/Doña	-0.374*** (0.016)	-0.337*** (0.021)	-0.176*** (0.022)	12,872 / 6,650
Interracial marriages	0.113*** (0.018)	0.113*** (0.018)	0.103*** (0.018)	2,869 / 2,869
Journeyman	0.211*** (0.019)	0.195*** (0.027)	0.130*** (0.029)	5,057 / 2,443
Laborer	0.171*** (0.017)	0.129*** (0.025)	0.089*** (0.026)	5,057 / 2,443
Similar occupations in hh.	0.034 (0.029)	-0.054 (0.040)	-0.071* (0.043)	2,154 / 1,230
Migrant	0.240*** (0.016)	0.073*** (0.026)	0.088*** (0.027)	13,376 / 6,662

Notes: East-minus-west differences in the 1821 *padrón*, on the sixteen *cuarteles* adjacent to the SJdD creek, for the outcomes of Figure A.11 that are not themselves measures of ethnicity. All columns control for sex and age indicators, with standard errors clustered by household in parentheses. Column (1) uses the full sample and reproduces the differences annotated in Figure A.11. Self-reported *calidad* is missing for about half the records, so column (2) re-estimates column (1) on the subsample in which it is recorded, and column (3) adds *calidad* to that same subsample. Columns (2) and (3) are therefore the comparison that isolates the control from the change in sample. The observations column reports the count for column (1) and for columns (2) and (3). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Figure A.12: Skin tone across the Guadalajara metropolitan area



(a) Distribution of skin tone

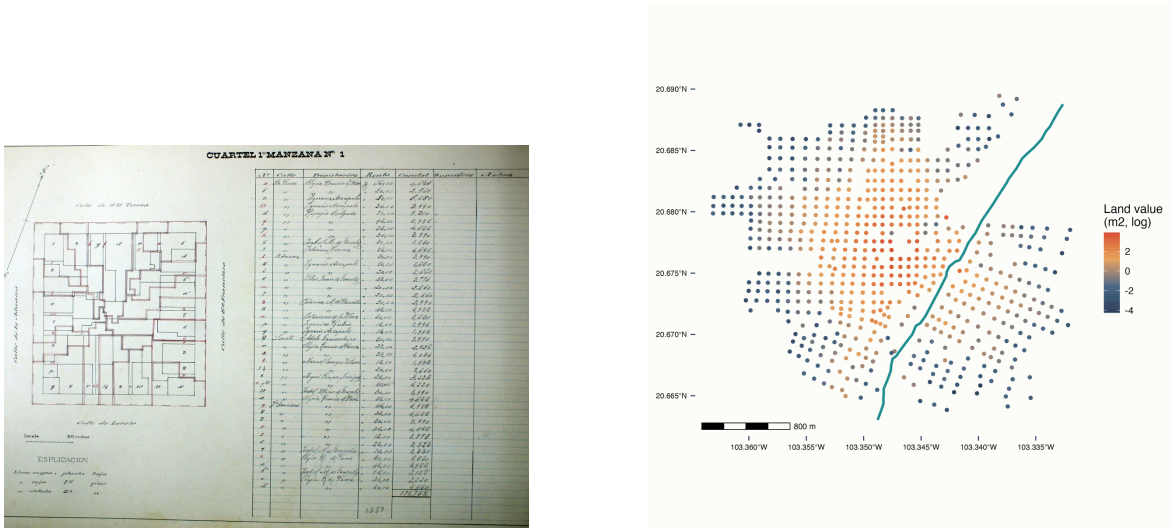
(b) Location of respondents

Notes: Interviewer-rated skin tone on the PERLA scale for respondents of the 2017 ESRU Survey of Social Mobility residing in the Guadalajara metropolitan area. Panel (a) plots the distribution of ratings; panel (b) locates each respondent, coloured by rating, with the SJdD boundary overlaid in blue. Ratings concentrate between 4 and 6 and are dispersed across the metropolitan area rather than clustering on either bank. The survey is not dense enough near the boundary to support a discontinuity estimate, so the figure describes the metropolitan pattern rather than testing balance at the cutoff.

A.8 Capitalization in the colonial core: 1882 land values

I introduce newly digitized microdata from Guadalajara's 1882 cadaster, a parcel-level record of assessed property values rare in its detail for nineteenth-century Latin America. Figure A.13, panel (a) reproduces an example façade entry from the original records; panel (b) shows the digitized distribution of land values across the central core. Historical maps were georeferenced to modern coordinates following the procedure detailed at the end of this subsection. Figure A.14

Figure A.13: Historical land values in 1882



(a) Property façade from the 1882 cadaster

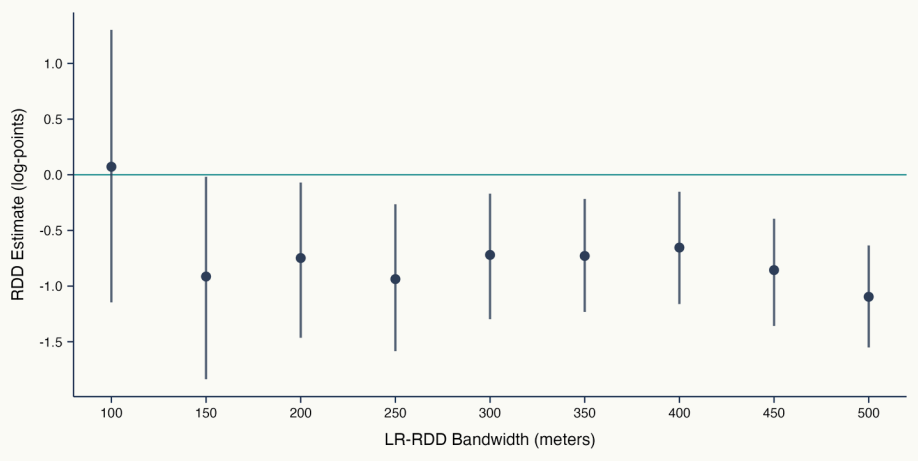
(b) Digitized land-value map (1882)

Notes: Panel (a) shows an example of a property façade entry as recorded in the 1882 cadaster of Guadalajara, alongside the formal assessed land value and an identifier locating the parcel within a city block (*manzana*). Panel (b) shows the digitized parcel-level land values across the historical core, with darker shades corresponding to higher assessed values per square meter. Historical cadastral maps were georeferenced to modern coordinates as described at the end of this subsection; the SJdD creek and the foundational settlements are overlaid. The clear west–east drop in assessed values across the creek is consistent with the formal RDD estimates in Figure A.14.

reports formal estimates of the discontinuity in 1882 land values at the SJdD creek using the Local Randomization RDD of Cattaneo et al. (2015). Across bandwidths from 150 to 500 meters, assessed land values are between 48 and 67 percent lower on the eastern (Analco) bank. The narrowest window, 100 meters, gives a positive point estimate with an interval too wide to be informative. This divergence in 1882 — well before 20th-century industrialization, the federal expansion of public schooling, or post-revolutionary demographic homogenization — captures the early capitalization of differential access to amenities, public goods, and social status into urban land markets, consistent with Baldomero-Quintana et al. (2025)'s evidence on land-value penalties from colonial Indigenous exclusion across Mexican cities.

Archival sources. The remainder of this subsection documents the construction of the land-value dataset underlying these estimates. The 1882 cadaster of Guadalajara is preserved at the *Archivo Municipal de Guadalajara* (AMG) in bound cadastral books organized by city block

Figure A.14: Local-randomization RDD estimates of land-value discontinuities, 1882



Notes: Each point reports the estimated discontinuity in log land values per square meter across the SJdD creek in 1882, using the Local Randomization RDD (LR-RDD) of Cattaneo et al. (2015), which treats observations within a small window around the boundary as if assigned at random. Estimates are reported across bandwidths from 100 to 500 meters; vertical bars show 95% confidence intervals. Negative values indicate lower assessed land prices on the historically excluded (eastern, Analco) side. Estimates are negative and statistically significant at every window from 150 meters outward, with point estimates between −0.65 and −1.10 log points, which corresponds to land values 48 to 67 percent lower on the eastern bank. The narrowest window, 100 meters, gives a positive point estimate whose interval runs from −1.15 to 1.30 and is uninformative. The cadastral data and georeferencing procedure are documented at the end of this subsection.

(*manzana*). Each *manzana* record contains, for every parcel, the property's unique identifier, street and house number, owner and renter names, structure type (*finca* versus *solar*), assessed land and structure values in pesos, and a free-text description often including a hand-drawn façade sketch and dimensions in *varas* (≈ 0.838 m). To my knowledge, this is the first time the 1882 cadaster has been systematically digitized and analyzed. The unit of analysis is the block rather than the parcel. For each *manzana* I add up the assessed land values of its *fincas* and divide the total by the surface of the block in square meters, which I obtain from the scale of the georeferenced cadastral map. The parcel dimensions recorded in *varas* do not enter this calculation, so the outcome does not depend on the legibility or completeness of the individual façade sketches, and no parcel is dropped for lacking a usable measurement. A block enters the estimation sample when its cadastral record and its polygon on the georeferenced map can both be identified.

Digitization. The cadastral books were photographed at the AMG and transcribed manually by a team of research assistants, with a first transcription pass and an independent verification pass; discrepancies were adjudicated against the original document, and illegible entries were flagged rather than imputed.

Georeferencing. Modern parcel-level geolocation is not directly available for the 1882 cadaster. I follow a three-step procedure. First, I obtain a high-resolution scan of the 1888 cadastral map published by the *Ayuntamiento de Guadalajara*, which uses the same *manzana* numbering. Second, I georeference the 1888 map to modern coordinates in QGIS using approximately 60 ground control points at locations unambiguously identifiable in both the historical map and modern satellite imagery (the Plaza de Armas, the Cathedral and surviving colonial churches, the *Mercado Corona*, and stable street intersections), using a thin-plate-spline transformation; the RMSE of the residuals is below 8 meters. Third, I match each 1882 parcel to its location in the georeferenced 1888 map by joining on the *manzana* number and within-manzana parcel ordering, prioritizing the 1882 record for substantive content. Parcels that cannot be unambiguously located account for less than 4% of the digitized parcels.

Validation. I cross-check parcel locations against (i) the modern *Catastro Municipal de Guadalajara*, (ii) Google Maps satellite imagery, and (iii) a sample of ≈ 200 parcels whose owner names can be matched to historical archives of notable families. Discrepancies arise primarily in zones of major 20th-century transformation, the construction of Calzada Independencia between 1903 and 1911 and the 1950s market expansions. No block is excluded from the estimation sample on this basis; every located block within a given window enters the corresponding LR-RDD.

Final dataset. The resulting dataset contains block-level records for the central core of Guadalajara, geolocated to modern coordinates, with assessed 1882 land value per m^2, structure value, ownership and use information, and the block's signed orthogonal distance to the historical SJdD creek. The land-value variable used in Figure A.14 is the log of assessed land value per m^2 at the block level. Replication code and the digitized dataset will be deposited publicly upon publication, conditional on AMG authorization.

A.9 Named institutions on each bank, 1800–1908

This subsection documents the evidence behind Section 4.1 and Figure 5: where the institutions named on the city's general maps stood relative to the SJdD creek, and how that compares with the population living on each bank.

Sources. Four general maps of Guadalajara carry an index that names the city's institutions and keys each one to a block. The 1800 sheet is a facsimile of a plan of the city "as it stood in the year 1800," copied from the original held at the *Archivo del Instituto de Geografía y Estadística*; its legend has 37 entries. The 1863 wave is a series of nine *cuartel* sheets held at the *Archivo Histórico de Jalisco*, surveyed by order of the state government for the use of its offices; the sheets for *cuarteles* 7 and 8 are signed by C. Vicente Ortigosa, 19 July 1863, and those for *cuarteles* 1 and 2 by the *Dirección General de Rentas*, 5 April 1863. Each sheet numbers its

blocks individually, labels its streets, and carries a legend of notable buildings. The 1890 wave is the *Plano General de la Ciudad de Guadalajara*, whose margins carry an index of establishments with *cuartel*, *manzana*, and street columns, divided into sections for notable buildings, markets, gardens, squares, arcades, hotels, baths, and former customs gates. The 1908 wave is the *Novísimo Plano General de la Ciudad de Guadalajara, con la nomenclatura últimamente aprobada por el Ayuntamiento*, by the engineer Regino Guzmán, published by Loreto y Ancira; it prints a red number on each block and lists the corresponding establishments in the margin under 105 entries.

Locating the entries. The 1863 and 1890 indices key their entries to a (*cuartel*, *manzana*) pair. I join them to the digitized 1890 plan, whose 713 polygons resolve into 672 numbered blocks once split pieces of the same block are merged, and whose numbering scheme the 1863 sheets share; the correspondence is validated against the 2010 national register of churches, restricted to within three kilometers of the creek before any name matching, with residuals between 0 and 75 meters. The 1800 legend has no block key: church entries are located through the same church register, and the remaining entries through name correspondence with the 1863 index. The 1908 plan requires reading the red numbers directly off the raster, which I do on magnified tiles; entries that survive in the earlier waves are additionally assigned by name correspondence with the georeferenced 1863 and 1890 sets. Every located entry is assigned to a bank by the sign of its position relative to the same boundary line used in the spatial RDD.

Table A.3 counts 31 of the 37 entries in the 1800 legend, all 56 in the 1863 sheets, 83 of the 91 in the 1890 index, and 62 of the 105 in the 1908 index. The six 1800 entries that do not enter are of four kinds. One, *el río*, names the creek itself rather than an establishment. One, *las garitas*, names the customs gates at several points of the perimeter and belongs to no single bank. The *Casa de Enseñanza de Niñas* could not be found after a directed search. The remaining three, the *Casa de Recogidas*, the *Real Fábrica de Tabacos* and the *Puente de las Damas*, are named on the sheet but could not be tied to a block precisely enough to place them, and the count admits only entries with a location. The eight 1890 entries that do not enter are the *Escuela de Jurisprudencia*, the *Hospital de la Trinidad*, the *Teatro Novedades*, the *Hipódromo*, the *Estación del FCCM*, the *Penitenciaría*, the *Jardín de la Plaza de Armas* and the *Jardín de Aclimatación*, whose index keys do not match a digitized block; several of them stood outside the surveyed grid. Of the 62 entries assigned in 1908, 24 were located on the plan and 38 by name correspondence with the georeferenced 1863 and 1890 sets. The 24 divide further into seventeen read directly off the raster, four placed through the church register, one the *Seminario Conciliar*, whose number the plan prints on two blocks, one on each bank, resolved to the western occurrence because it matches its 1890 location, and two assigned by elimination after failing to appear in eighteen sampled tiles of the central grid. Elimination is a way of locating an entry on the plan rather than a third method beside the other two. The educational series is counted from a second and more intensive pass over the sheets, which admits entries read directly off the plan as well as georeferenced ones, and therefore recovers entries the general count leaves out. Among the fifteen educational entries

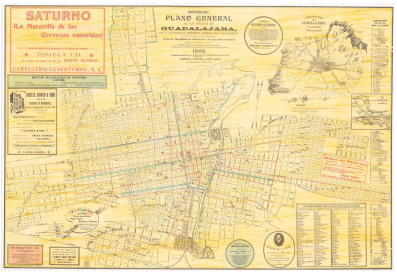
Figure A.15: Cartographic sources for the count of named institutions, 1800–1908



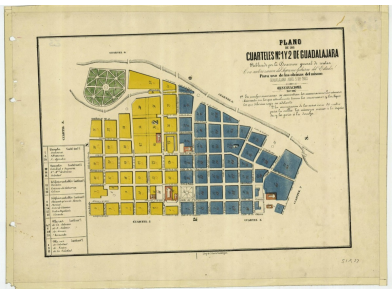
(a) 1800 facsimile plan



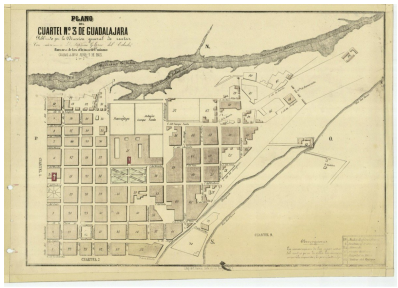
(b) 1890 *Plano General*



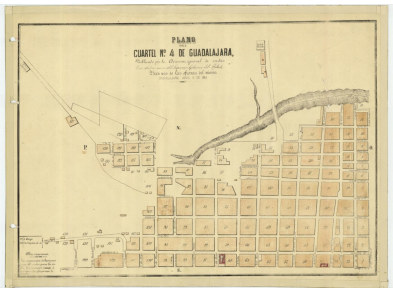
(c) 1908 *Novísimo Plano General*



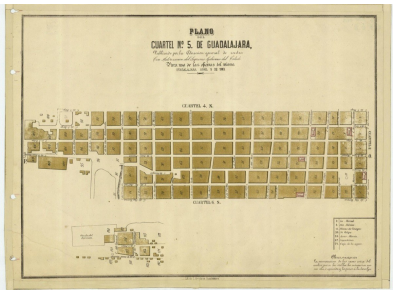
(d) 1863, *cuarteles* 1 and 2



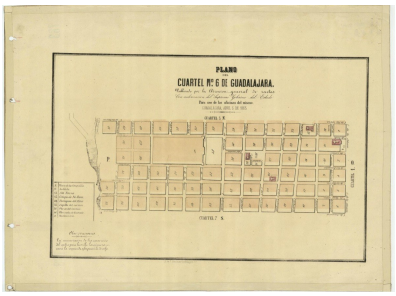
(e) 1863, *cuartel* 3



(f) 1863, *cuartel* 4



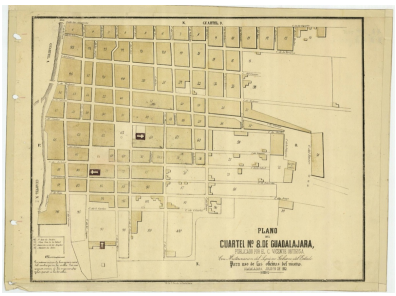
(g) 1863, *cuartel* 5



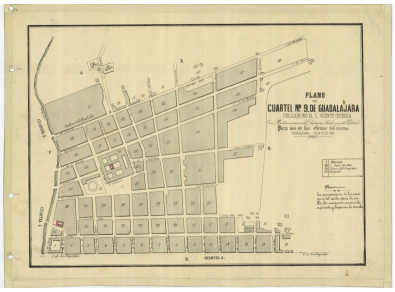
(h) 1863, *cuartel* 6



(i) 1863, *cuartel* 7



(j) 1863, *cuartel* 8 (Analco)



(k) 1863, *cuartel* 9

Notes: The four waves of Table A.3 and Figure 5. Panels (a), (b) and (c) are the three general plans, each carrying an index in its margins that names the city's institutions and keys them to a block. Panels (d) through (k) are the 1863 *cuartel* survey held at the *Archivo Histórico de Jalisco*, eight sheets covering the city's nine *cuarteles*; the sheet in panel (d) covers *cuarteles* 1 and 2 together. Each 1863 sheet numbers its blocks individually, labels its streets, and carries its own legend of notable buildings, which is what allows an entry to be keyed to a (*cuartel*, *manzana*) pair. *Cuartel* 8, in panel (j), is Analco, on the eastern bank of the SJdD creek. Images are reproduced at reduced resolution; the counting was done on the full-resolution scans.

of the 1908 index, fourteen are western and one eastern; of the fourteen, twelve were located directly and two by elimination. In 1800 the same pass recovers two western educational entries, the *Colegio femenino* and the *Iglesia de Ex-Jesuitas y Universidad*.

Counts by category. Table A.3 reports the located entries by category, wave, and bank.

Table A.3: Located institutions by category, wave, and bank

	1800		1863		1890		1908	
Category	West	East	West	East	West	East	West	East
Church	15	3	18	3	0	0	21	3
Educational	4	0	3	0	10	0	14	1
Government	5	0	5	0	12	1	5	0
Square and garden	0	0	11	1	8	0	6	0
Commerce	0	0	0	0	20	0	0	0
Culture and recreation	1	0	4	0	7	1	2	1
Health and charity	1	0	1	1	5	2	3	2
Market	0	0	2	1	3	1	3	1
Baths	0	0	0	0	8	1	0	0
Infrastructure	2	0	1	0	3	0	0	0
Mill	0	0	3	0	0	0	0	0
Cemetery	0	0	1	1	0	1	0	0
Total	28	3	49	7	76	7	54	8

Notes: Counts of entries in each map's index that could be located on the urban grid and assigned to a bank of the SJdD creek. Categories are assigned by the author from the entry names and the sections of the original indices. Each map reflects its own editorial conventions about what merits an entry, so the composition of categories varies across waves: the 1890 index has a commercial section and no separate church section, while the 1908 index lists churches and no commercial establishments. Levels are therefore comparable across banks within a wave, and across waves only within a category that appears in all four, as the educational category does.

In 1800, the only entries east of the creek are the three churches of Analco, San Sebastián, and San Juan de Dios. The eastern bank appears in the index of that year with places of worship and with no institution of any other kind.

Population on each bank. Two enumerations record residence by *cuartel* close enough to these waves to serve as a denominator. To assign *cuarteles* to banks I use the digitized blocks of the 1890 plan. *Cuartel* 9 lies entirely east of the creek and *cuarteles* 1 through 7 entirely west, with no block on the far side of the line in either case. *Cuartel* 8 has 87 of its 91 blocks east. Those four blocks are the only departure from a partition that is otherwise exact at the level of the *cuartel*, and they are too few to make any other assignment of *cuartel* 8 sensible.

In the 1821 *padrón* of Anderson and Spike (2007), the two eastern *cuarteles* hold 6,873 residents against 10,346 in the seven western *cuarteles* of the central core, an eastern share of 39.9 percent. The comparison is restricted to that core, which is also the area the plans cover, and accounts for 49 percent of the 35,287 residents enumerated in 1821. Of the 24 populated *cuarteles* of the *padrón*, the remaining fifteen lie outside the core and cannot be assigned to a bank from what the enumeration records. That assignment rests on the *padrón*'s *cuartel* numbers denoting the same units as the plans, and the *padrón* does not allow me to verify it. Its scheme does not correspond to the one the maps use, and the natural check, matching street names, is unavailable for the unit that matters most, since the street field is empty in all 3,892 records of *cuartel* 8, the most populous unit in the enumeration and the one taken to be Analco. The 1821 share is therefore a benchmark conditional on that identification rather than an independently verified denominator. The 1930 census does not carry the same ambiguity. Its *cuartel* scheme runs from 1 to 11 and appears to continue the one the plans use, and there the two eastern *cuarteles* hold 33.6 percent of the population recorded for the city.

Limitations. Three limitations bear on how these counts should be read. First, the indices record the institutions their authors considered worth naming, and neighborhood primary instruction would not ordinarily appear in them; the counts describe where the city's endowed establishments stood rather than the number of school places available on each bank. Second, the share of located entries in 1908 is 59 percent, and the excluded entries are concentrated in categories that expanded during the Porfiriato, so the eastern share in that wave may be overstated. Third, the transcriptions of the indices have been coded once; a second independent reading of a subsample, with an agreement rate, has not yet been carried out.

B Persistent Human Capital Gaps

This appendix collects the supporting evidence for the spatial RDD analysis in Section 3, in the order the section draws on it. Appendix B.1 reports the density test and Appendix B.2 the full set of main estimates and robustness specifications. Appendix B.3, Appendix B.4, and Appendix B.5 address housing-market sorting, school supply and quality, and crime exposure. Appendix B.6 estimates the discontinuity point by point along the boundary and Appendix B.7 maps the schooling gradient across the metropolitan area. Appendix B.8 and Appendix B.9 report additional multi-wave outcomes and the 1930 census comparison. Appendix B.10 reports the donut estimates and Appendix B.11 the placebo-boundary exercise. Appendix B.12 addresses residential turnover and selective in-migration, and Appendix B.13 details the design behind the 1992-explosions evidence.

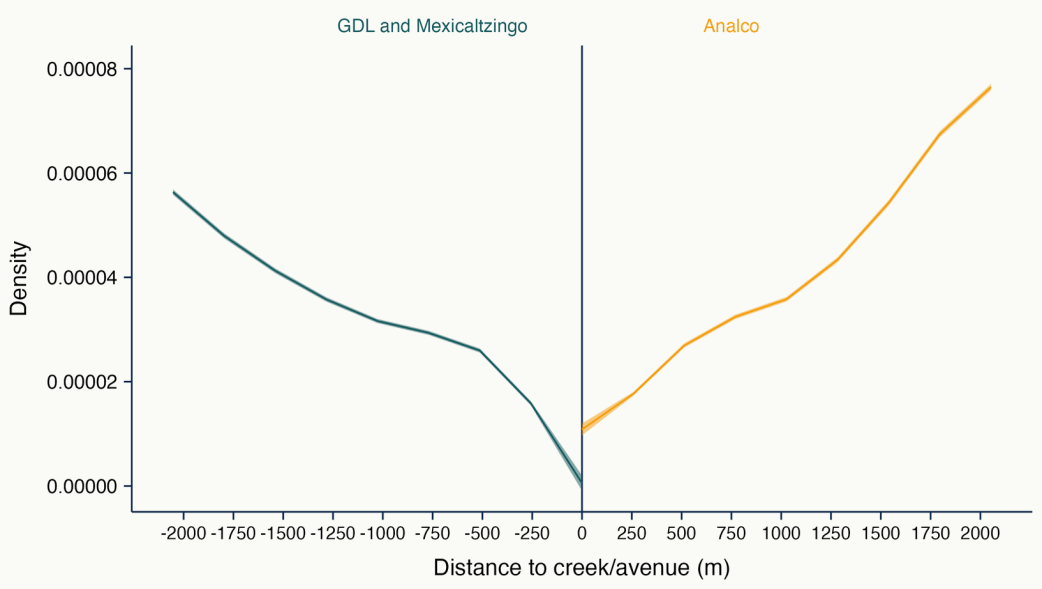
B.1 Density at the boundary

A first concern in spatial RDD designs is whether the density of observations is itself discontinuous at the boundary (McCrary, 2008; Cattaneo et al., 2018). Figure B.1 plots the estimated density of individuals as a function of distance to the boundary, applying Cattaneo et al. (2018). The density declines toward zero on both sides, reflecting the lower residential population in the historical core. At the cutoff itself the estimated density is higher on the Analco side, and the two banks converge to within four percent of each other by 500 meters. Density rises monotonically with distance into Analco rather than piling up at the line, which is the pattern that manipulation would produce. It does not rule out the gradual sorting that Section 3 addresses with cadastral land values and measures of residential persistence, and its power is limited where the density is lowest.

B.2 Main estimates and full robustness battery

Table B.1 reports the full set of spatial RDD estimates underlying Figure 2. The table is organized in two panels — years of schooling for individuals aged 6+ (Panel A) and school attendance for ages 6–22 (Panel B) — with, for each outcome, four main specifications followed by robustness checks. The four main specifications progress from the unconditional levels estimate to the standardized estimate, add clean covariates, and finally add controls that may be partially affected by the treatment ("unclean" controls whose inclusion risks collider bias). Each robustness check holds the clean controls fixed and varies one design choice: the local-polynomial order, the kernel, the variance estimator, or the sample. Across the standardized specifications, the boundary effect on years of schooling lies between −0.08 and −0.17 SD, centered on the preferred covariate-adjusted estimate of −0.14. School-attendance estimates are noisier but consistently negative (−0.08 to −0.14 SD). The Calonico–Cattaneo–Titiunik nearest-neighbor (NN) variance estimator

Figure B.1: Density test at the historical boundary



Notes: Local-polynomial density estimate of individuals as a function of signed orthogonal distance to the historical boundary (Calzada Independencia, formerly the SJdD creek), following Cattaneo et al. (2018). The horizontal axis reports distance in meters, with positive values on the eastern (Analco) side. Shaded bands report 95% confidence intervals. Density declines toward zero on both sides, reflecting the lower residential population in the historical core. The estimated density at the cutoff is higher on the eastern side, and the two sides converge to within four percent by 500 meters. The absence of a spike at the line is evidence against bunching; it is not evidence against gradual sorting across the two banks, and the test has limited power where the density approaches zero.

is used in the main specifications; the HC3 row reports the most conservative heteroskedasticity-consistent alternative.

Table B.1: Spatial RDD estimates on human capital, 2020 Census

$\hat{\tau}$	SE	95% CI	p-val	Poly	h	N_h	b	N_b	Kernel	VCE	Clean ctrls	Unclean ctrls	Specification
Panel A. Years of schooling (ages 6+)													
-0.791	0.156	[-1.097, -0.485]	0.000	1	570.750	25011	1489.073	107104	Uniform	NN	No	No	Main – levels
-0.165	0.033	[-0.229, -0.101]	0.000	1	570.750	25011	1489.073	107104	Uniform	NN	No	No	Main – z-score
-0.141	0.048	[-0.235, -0.047]	0.003	1	395.465	13833	769.946	42070	Uniform	NN	✓	No	Main – z-score + covariates
-0.117	0.041	[-0.198, -0.036]	0.005	1	410.825	13465	796.536	40669	Uniform	NN	✓	✓	Main – z-score + covariates + clean controls
-0.151	0.047	[-0.244, -0.058]	0.001	2	741.235	39262	1461.118	104303	Uniform	NN	✓	No	Rob. – Polynomial order (2)
-0.103	0.047	[-0.194, -0.012]	0.027	1	477.508	18525	897.343	51626	Triangular	NN	✓	No	Rob. – Kernel (triangular)
-0.127	0.044	[-0.212, -0.042]	0.004	1	504.318	20665	913.445	53039	Epanechnikov	NN	✓	No	Rob. – Kernel (Epanechnikov)
-0.081	0.051	[-0.181, 0.018]	0.110	1	388.115	13627	759.784	41158	Uniform	HC3	✓	No	Rob. – VCE (HC3, most conservative)
-0.122	0.043	[-0.205, -0.039]	0.004	1	407.383	12717	796.109	38784	Uniform	NN	✓	✓	Rob. – Migrants (native-born only)
Panel B. School attendance (ages 6–22)													
-0.032	0.019	[-0.070, 0.005]	0.094	1	1200.216	18265	2058.598	45016	Uniform	NN	No	No	Main – levels
-0.075	0.045	[-0.163, 0.013]	0.094	1	1200.216	18265	2058.598	45016	Uniform	NN	No	No	Main – z-score
-0.104	0.072	[-0.245, 0.036]	0.146	1	571.045	6101	1081.865	15789	Uniform	NN	✓	No	Main – z-score + covariates
-0.086	0.047	[-0.178, 0.006]	0.067	1	894.781	12172	1997.570	41693	Uniform	NN	✓	✓	Main – z-score + covariates + clean controls
-0.129	0.082	[-0.289, 0.032]	0.116	2	871.024	12038	1800.565	35228	Uniform	NN	✓	No	Rob. – Polynomial order (2)
-0.136	0.063	[-0.260, -0.011]	0.032	1	703.139	8847	1618.948	29378	Triangular	NN	✓	No	Rob. – Kernel (triangular)
-0.124	0.068	[-0.257, 0.009]	0.067	1	616.138	7038	1472.523	24574	Epanechnikov	NN	✓	No	Rob. – Kernel (Epanechnikov)
-0.091	0.078	[-0.244, 0.063]	0.246	1	504.945	5117	1028.982	14819	Uniform	HC3	✓	No	Rob. – VCE (HC3, most conservative)

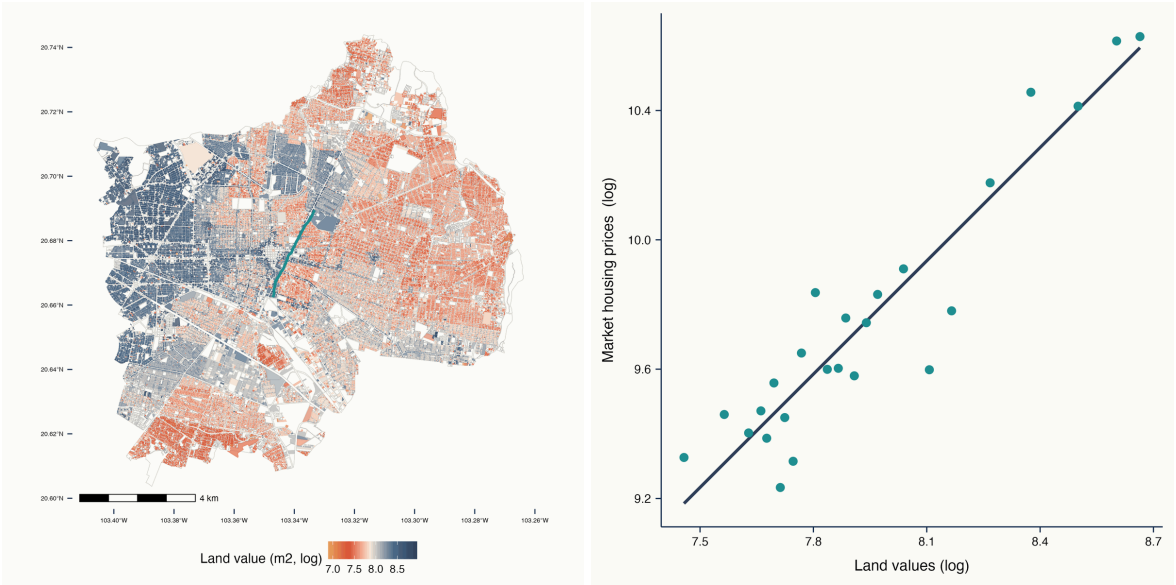
Notes: Spatial RDD estimates of the boundary effect on human capital outcomes, using individual-level data from the complete 2020 *Censo de Población y Vivienda* for the municipality of Guadalajara. The running variable is the signed orthogonal distance from the resident's geolocated block to the historical SJdD creek (today Calzada Independencia), with positive values on the eastern (Analco) side. Inference follows Calonico et al. (2014): $\hat{\tau}$ is the bias-corrected treatment effect; SE the robust bias-corrected standard error; h and b the MSE-optimal main and bias-correction bandwidths (meters); N_h and N_b the effective sample sizes. *Poly* is the local polynomial order; *Kernel* the weighting scheme; *VCE* the variance estimator (NN the Calonico–Cattaneo–Titiunik nearest-neighbor estimator; HC0–HC3 the heteroskedasticity-consistent estimators). *Clean controls* include sex, age, longitude, latitude, and boundary-segment fixed effects. *Unclean controls* in Panel A include indicators for Afrodescendant or Indigenous self-identification, birth in Jalisco, residence in the municipality five years earlier, social security access, marital status, and dwelling type — variables that may themselves be affected by treatment. Unclean controls in Panel B include parental mean years of schooling, household size, and residence in the municipality five years earlier. The preferred estimate is the third row in Panel A and the fourth in Panel B. The two panels differ because the attendance comparison in Section 3 is between children of comparable households, and parental schooling and household size are the inherited human-capital stock that comparison holds fixed, at the cost of conditioning on variables measured after the assignment. The third row of Panel B, which omits them, gives −0.104 with a conventional p-value of 0.146 against 0.067 for the fourth. "Main – levels" coefficients are in original units (years of schooling for Panel A); all others are in standardized (z-score) units.

B.3 Housing-market sorting and capitalization

If housing prices capitalize amenities or the boundary's stigma, socioeconomic groups may sort across space, and gaps could reflect selection rather than the legacy of exclusion.

The west–east gradient in land values. Figure B.2a maps parcel-level cadastral land values, documenting a clear west–east gradient. Figure B.2b validates the cadastral measure against modern market housing prices at the ZIP-code level (IIEG web-scraped listings); the two are tightly correlated, so the cadastral measure is a credible proxy for the price-relevant component of land value.

Figure B.2: Land values and housing-market capitalization in Guadalajara



(a) Spatial distribution of land values (log) across Guadalajara

(b) Cadastral land values vs. market housing prices (log)

Notes: Panel (a) maps parcel-level cadastral land values (log pesos per m^2) from the Guadalajara Municipal Cadaster (2023). Each polygon is a parcel; shading runs from red (lower) to blue (higher). The thick green line marks Calzada Independencia. Parcels east of the boundary are systematically less valuable per m^2. Panel (b) validates the cadastral measure against modern market housing prices. Both variables are aggregated to the neighborhood (*colonia*): the horizontal axis is the log of the neighborhood mean of cadastral land value per m^2 and the vertical axis the log of the neighborhood mean of market housing price per m^2 from web-scraped listings, so the two panels are on the same logarithmic scale. Each point is one of 25 bins; the solid line is a linear fit and the dashed line is the 45-degree line. Sources: Guadalajara Municipal Cadaster (2023); IIEG web-scraped housing listings.

Robustness of the boundary effect to land-value controls. Table B.2 reports spatial RDD estimates after adding flexible controls for block-level land values. The preferred specification adds land values to the clean controls, giving −0.39 standardized block-level units, or about 0.78 years of schooling. Across the land-value specifications the estimate ranges from −0.39 to −0.49 in block-level standard deviations, that is from 0.78 to 0.97 years, which brackets the estimate obtained on individual data; land values absorb part of the cross-boundary differential without changing the magnitude the individual-level design reports.

Table B.2: Spatial and geographic RDD estimates controlling for land values (block level, 2020 Census)

Specification	$\hat{\tau}$	SE	95% CI	p-val	Poly	h	b	N_h
Main (unweighted)	-0.622	0.189	[-0.992, -0.253]	0.001	1	821.3	1651.0	593
Main (Poly 2)	-0.609	0.175	[-0.952, -0.267]	0.000	2	1820.7	2859.2	1,797
With clean controls	-0.561	0.176	[-0.906, -0.215]	0.001	1	698.1	1536.1	470
With clean controls (Poly 2)	-0.641	0.179	[-0.993, -0.290]	0.000	2	1468.8	2629.8	1,309
+ Land value control	-0.391	0.150	[-0.685, -0.098]	0.009	1	853.8	1706.7	622
+ Land value control (Poly 2)	-0.486	0.156	[-0.793, -0.180]	0.002	2	1730.6	2692.5	1,614

Notes: Spatial RDD estimates of the boundary effect on mean years of schooling at the census-block level, using aggregated 2020 census data for Guadalajara. Coefficients are in standardized (z-score) units; block-level mean years of schooling is 11.22 (SD 2.00). "Main (unweighted)" is the baseline block-level RDD without land-value controls. "Clean controls" include block longitude, latitude, neighborhood fixed effects, and boundary-segment dummies. "Land value control" adds the log of cadastral land values (pesos per m^2). All specifications are unweighted, so that the census block is the unit of analysis; the person-weighted estimand is the individual-level estimate reported in Table B.1. Inference follows Calonico et al. (2014); h and b are the MSE-optimal main and bias-correction bandwidths (meters); N_h the effective sample size within the main bandwidth. The preferred specification is "+ Land value control": clean controls and land values. Land-value controls leave the effect economically and statistically significant, and equal to between 0.78 and 0.97 years once block-level standardized units are converted at the block standard deviation of 2.00.

Boundary-adaptive estimates with land-value controls. Figure B.3 extends the boundary-adaptive (bivariate) RDD of Cattaneo et al. (2025) to the land-value-controlled specification. The pointwise estimates remain negative and concentrated near the historical core.

B.4 School supply and quality

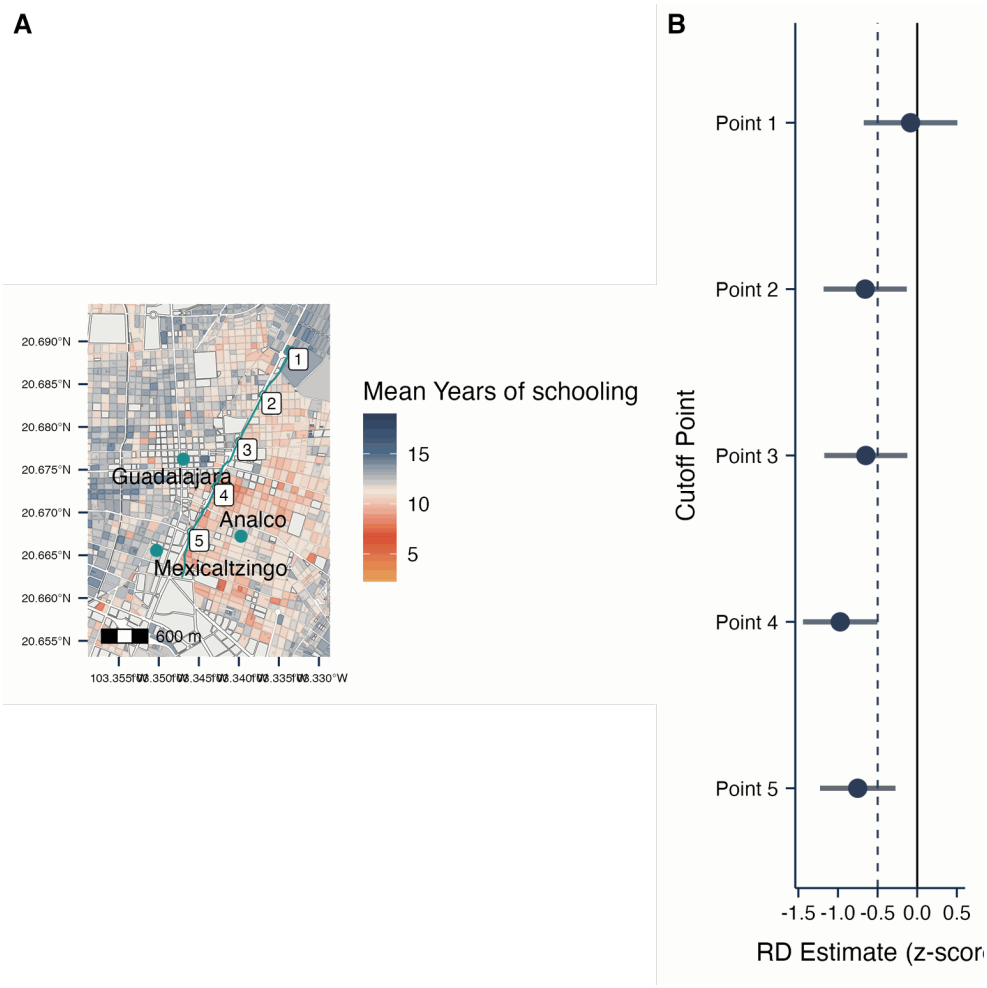
School supply. Table B.3 reports spatial RDD estimates of the boundary effect on the log number of schools per block, by sector and level, from 2018–2019 administrative records. None of the eight outcomes shows a significant discontinuity; all q-values are at unity after FDR adjustment. School supply is balanced across the boundary on every dimension.

School quality. I assess quality via *Formato 911* (approval, repetition, dropout, class size, student–teacher ratios; 1998–2015) and ENLACE standardized tests (Spanish and Math; 2006–2014). For each source I run a municipality-wide OLS comparison (Figures B.4 and B.6) and a spatial RDD within 1.5 km of the boundary (Figures B.5 and B.7). All four exercises show no systematic quality differential across the boundary. School access and quality are not the channel.

B.5 Crime exposure

Figure B.8 maps reported crimes (*carpetas de investigación*) in central Guadalajara (2017–2024). Table B.4 formalizes the comparison: none of the ten estimates is significant after FDR adjustment, and the largest unadjusted estimates (violence against women) have q-values around

Figure B.3: Boundary-adaptive (geographic) RDD with block-level land-value controls



Notes: Boundary-adaptive bivariate RDD estimates and 95% uniform confidence bands following Cattaneo et al. (2025), using block-level 2020 census data for Guadalajara and including block-level controls for log cadastral land values. The horizontal axis indexes points along the boundary. The shaded region is the 95% uniform confidence band; the dashed line marks the pooled (AATE) estimate. The vertical axis is in standardized units of mean years of schooling (sample mean 11.22, SD 2.00). Pointwise estimates remain negative along the boundary, consistent with the univariate land-value-controlled estimate in Table B.2.

0.7. The boundary does not coincide with a discontinuity in crime exposure.

B.6 Estimates along the boundary

Because the historical boundary is irregular, Figure 3 estimates the discontinuity at each point along it. Table B.5 reports those estimates. The two southernmost points, which span the segment facing Mexicaltzingo, are the only ones whose uniform confidence band lies entirely below zero in both outcomes, and each differs significantly from every northern point. The two southern points do not differ from each other, and the northern points do not differ among themselves.

B.7 The east–west pattern beyond the municipality

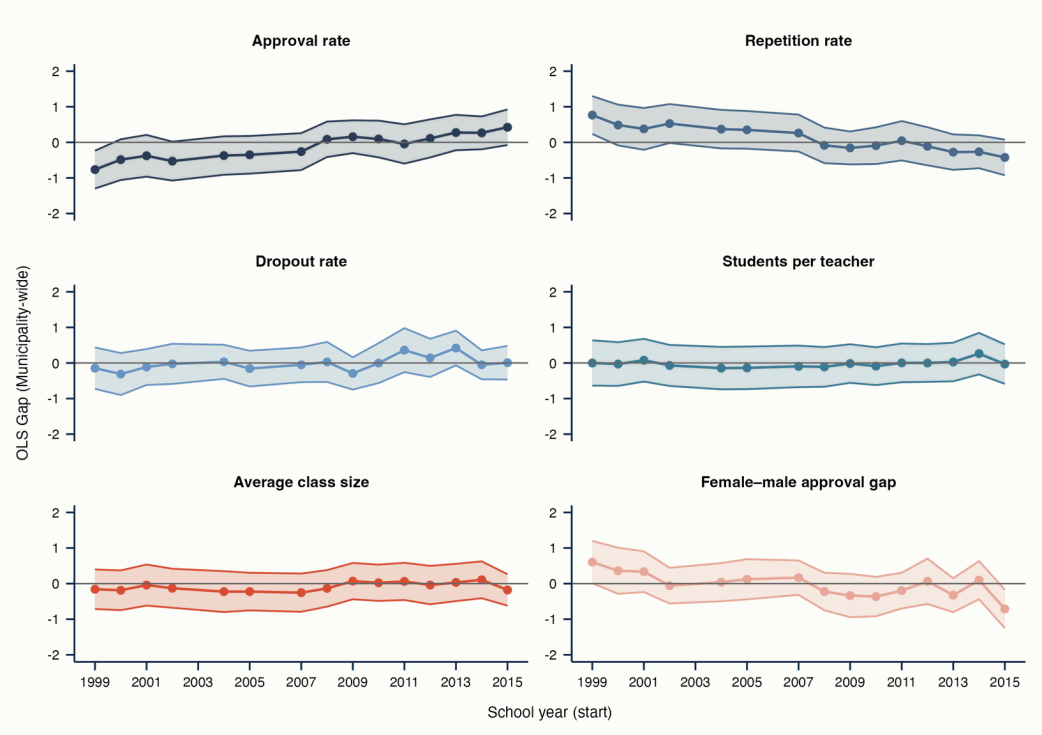
Figure B.9 extends the block-level map of Figure 1b to the ten municipalities of the Guadalajara metropolitan area. The east–west gradient does not stop at the municipal border: blocks west

Table B.3: Spatial RDD estimates on school supply (block level, 2018–2019)

Outcome	$\hat{\tau}$	SE	95% CI	p-val	q-val	Poly	h	b	N_h
Total schools (log)	0.128	0.177	[-0.218, 0.475]	0.468	1.000	1	1574.3	2634.0	278
Total schools × shifts (log)	0.006	0.210	[-0.407, 0.418]	0.979	1.000	1	1545.2	2644.5	278
Public schools (log)	0.289	0.307	[-0.312, 0.891]	0.346	1.000	1	1360.6	2339.4	232
Private schools (log)	-0.181	0.370	[-0.906, 0.544]	0.625	1.000	1	1644.4	2737.7	296
Primary schools (log)	0.221	0.243	[-0.256, 0.698]	0.363	1.000	1	1624.9	2684.2	285
Secondary schools (log)	-0.074	0.116	[-0.301, 0.154]	0.526	1.000	1	1826.5	3376.2	392
High schools (log)	-0.348	0.234	[-0.806, 0.111]	0.137	1.000	1	2304.2	3791.6	451
Public share	0.273	0.284	[-0.284, 0.830]	0.337	1.000	1	1409.6	2382.0	240

Notes: Spatial RDD estimates of the boundary effect on educational infrastructure, using block-level data for Guadalajara from administrative school records for 2018–2019. Dependent variables: log total schools per block; log school shifts (*turnos*); logs of public-only, private-only, primary, secondary, and high-school counts; and the public share. All log outcomes use $\log(1+x)$. Inference follows Calonico et al. (2014); specifications include boundary-segment fixed effects and geographic coordinates. The q-value column reports sharpened FDR-adjusted p-values across the eight outcomes (Anderson, 2008); all are at unity. School supply is balanced on every dimension.

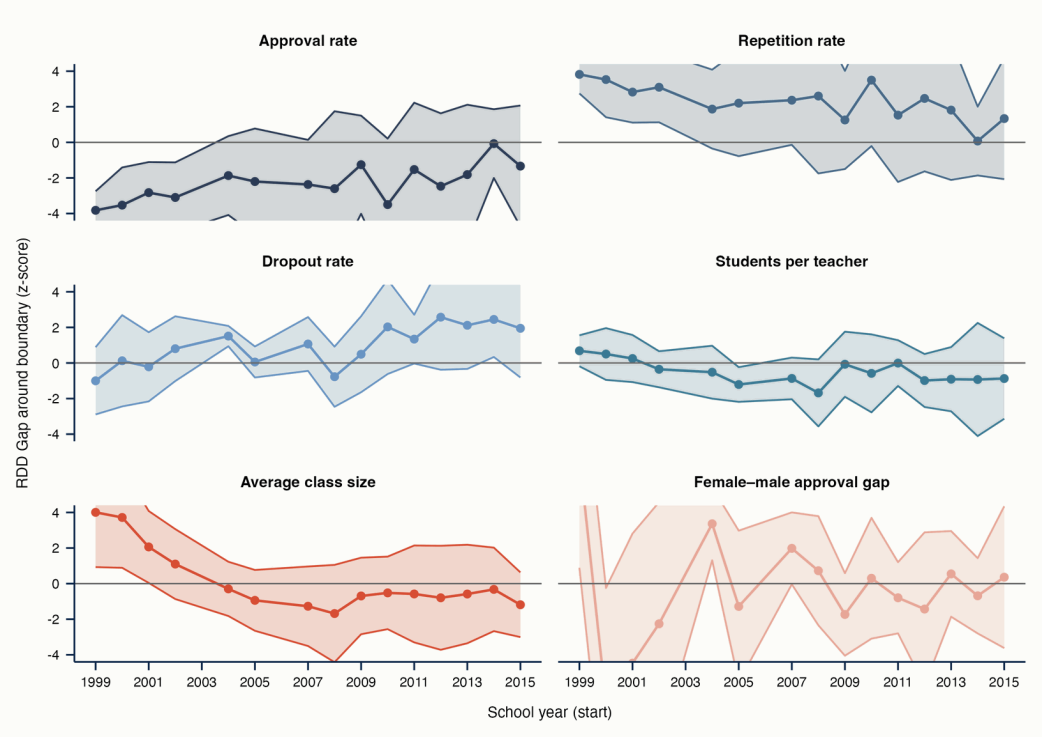
Figure B.4: Municipality-wide OLS estimates of school quality (Formato 911, secondary)



Notes: East-minus-west gaps in school-level operational indicators for secondary schools across Guadalajara, using *Formato 911* data (1998–2015). Each panel reports a separate outcome (approval, repetition, dropout, class size, student–teacher ratio, female-minus-male approval gap), standardized within year. Specifications include school-location controls, a public/private indicator, and year fixed effects. Confidence bands are 95% and are pointwise, without adjustment for the six outcomes or the seventeen years. The approval and repetition gaps favour the western schools at the start of the series, with bands away from zero through the early 2000s, and close by the middle of the decade. The remaining indicators fluctuate around zero throughout.

of the historical core, into Zapopan, hold the highest mean schooling, and blocks east of it, into Tonalá and San Pedro Tlaquepaque, the lowest.

Figure B.5: Boundary RDD estimates of school quality (Formato 911, secondary)



Notes: Spatial RDD estimates of the boundary effect on school-level operational indicators for secondary schools within 1.5 km of the boundary, *Formato 911* (1998–2015), standardized within year, following Calonico et al. (2014). Confidence bands are 95% and are pointwise, without adjustment for the six outcomes or the seventeen years. The approval, repetition and class-size gaps favour the western schools at the start of the series, with bands away from zero through about 2002, and close over the following few years. From the mid-2000s onward none of the six indicators differs systematically from zero.

Table B.4: Spatial RDD estimates on crime opportunities (block level, 2017–2024)

Outcome	$\hat{\tau}$	SE	95% CI	p-val	q-val	Poly	h	b	N_h
Panel A. Total crime									
Log density	0.151	0.247	[-0.333, 0.635]	0.541	1.000	1	766.0	1201.8	1,297
Log rate	0.094	0.238	[-0.374, 0.561]	0.694	1.000	1	867.9	1261.4	1,382
Panel B. By crime type									
Property crimes (log density)	0.024	0.200	[-0.368, 0.416]	0.904	1.000	1	714.5	1142.7	1,217
Property crimes (log rate)	-0.030	0.204	[-0.430, 0.371]	0.885	1.000	1	823.5	1196.8	1,291
Violent crimes (log density)	0.044	0.069	[-0.091, 0.179]	0.520	1.000	1	912.0	1415.6	1,604
Violent crimes (log rate)	0.021	0.067	[-0.110, 0.153]	0.749	1.000	1	1152.9	1727.4	2,062
Homicide (log density)	0.011	0.019	[-0.026, 0.048]	0.560	1.000	1	735.2	1324.4	1,466
Homicide (log rate)	0.020	0.016	[-0.012, 0.051]	0.214	1.000	1	969.4	1495.3	1,707
Violence Against Women (log density)	0.148	0.086	[-0.020, 0.317]	0.084	0.725	1	831.5	1287.4	1,416
Violence Against Women (log rate)	0.123	0.070	[-0.015, 0.261]	0.081	0.725	1	884.2	1370.1	1,532

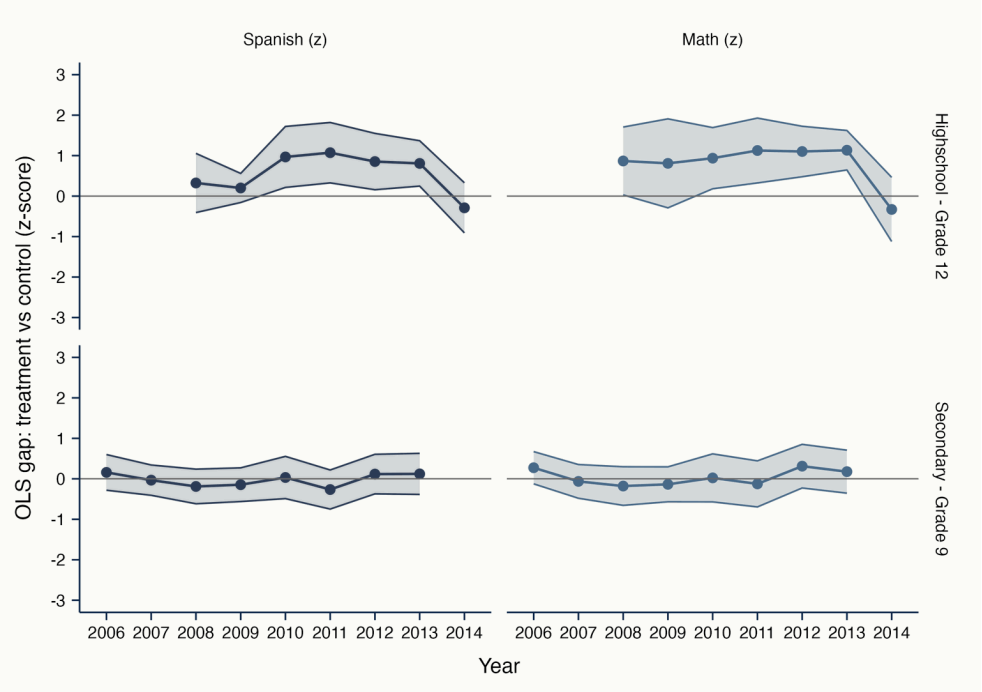
Notes: Spatial RDD estimates of the boundary effect on crime measures, block-level data for Guadalajara from the Jalisco Attorney General's Office (2017–2024). Crime density is incidents per km^2; crime rate is incidents per 1,000 residents. All outcomes are $\log(1 + x)$. "Property" includes theft, burglary, vehicular theft; "Violent" intentional injuries and assaults excluding homicide; "Homicide" intentional homicide; "VAW" domestic violence, sexual crimes, feminicide. Inference follows Calonico et al. (2014); specifications include boundary-segment fixed effects and geographic coordinates. The q-value column reports sharpened FDR-adjusted p-values across the ten outcomes (Anderson, 2008). None is significant after correction.

Table B.5: Boundary-adaptive RDD estimates by boundary point

	Years of schooling		School attendance	
	Estimate (s.e.)	95% uniform band	Estimate (s.e.)	95% uniform band
Panel A. Pointwise estimates along the boundary				
Point 1 — North	−0.427	[−1.60, 0.75]	0.073	[−0.02, 0.17]
	(0.457)		(0.038)	
Point 2 — North	−0.724	[−1.40, −0.04]	−0.033	[−0.12, 0.05]
	(0.264)		(0.034)	
Point 3 — North	−0.894	[−2.23, 0.45]	0.013	[−0.04, 0.06]
	(0.522)		(0.020)	
Point 4 — South (Mexicaltzingo)	−2.291	[−3.06, −1.52]	−0.126	[−0.20, −0.05]
	(0.299)		(0.030)	
Point 5 — South (Mexicaltzingo)	−1.945	[−2.66, −1.23]	−0.118	[−0.21, −0.03]
	(0.277)		(0.036)	
Panel B. Pairwise differences, southern segment minus each northern point				
	Difference (s.e.)	q-value	Difference (s.e.)	q-value
Point 4 − point 1	−1.863***	0.002	−0.198***	0.000
	(0.547)		(0.048)	
Point 4 − point 2	−1.567***	0.001	−0.092**	0.016
	(0.399)		(0.045)	
Point 4 − point 3	−1.397**	0.008	−0.139***	0.000
	(0.602)		(0.035)	
Point 5 − point 1	−1.517***	0.004	−0.190***	0.000
	(0.535)		(0.052)	
Point 5 − point 2	−1.221***	0.002	−0.084*	0.029
	(0.383)		(0.049)	
Point 5 − point 3	−1.051*	0.025	−0.131***	0.001
	(0.591)		(0.041)	

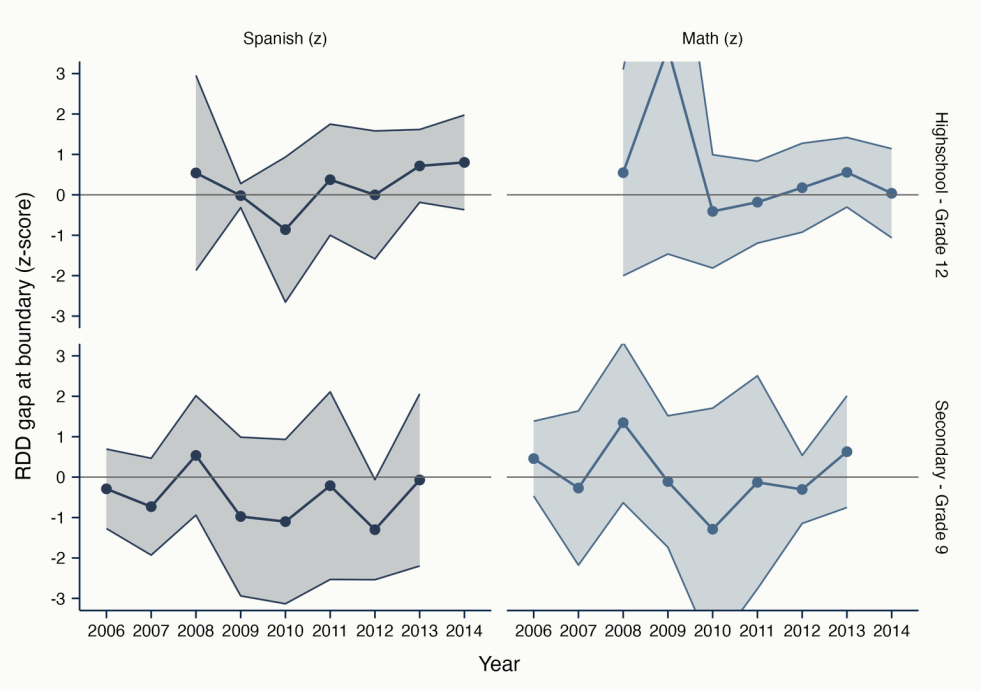
Notes: Boundary-adaptive RDD estimates following Cattaneo et al. (2025), using individual-level data from the 2020 *Censo de Población y Vivienda* geolocated at the block level. Boundary points are ordered from north to south; points 4 and 5 span the southern segment, the one facing the historical *pueblo* of Mexicaltzingo. Panel A reports the robust bias-corrected estimate with its standard error and the 95% uniform confidence band, which is centred on that estimate and is simultaneous over the five points reported. These are the estimates plotted in Figure 3, so the table and the figure report the same quantities. Panel B reports the difference between each southern point and each northern point, computed from the bias-corrected estimates. Because `rd2d` does not export the covariance between boundary points, and the version available when these estimates were produced offered no joint test of equality across points, the standard error of the difference treats the two estimates as independent. That overstates the standard error when the covariance between the two estimates is positive, which overlapping local neighbourhoods would ordinarily produce, but bias-corrected local-polynomial estimates carry influence weights that can change sign, so the direction is not

Figure B.6: Municipality-wide OLS estimates of standardized test performance (ENLACE)



Notes: East-minus-west gaps in mean ENLACE scores in Spanish and Math at the last grade of secondary and high school across Guadalajara (2006–2014). Specifications include school-location controls, a public/private indicator, and year fixed effects. Confidence bands are 95%. These scores begin in 2006, after the convergence visible in Figure B.4, and the gaps fluctuate around zero throughout.

Figure B.7: Boundary RDD estimates of standardized test performance (ENLACE)

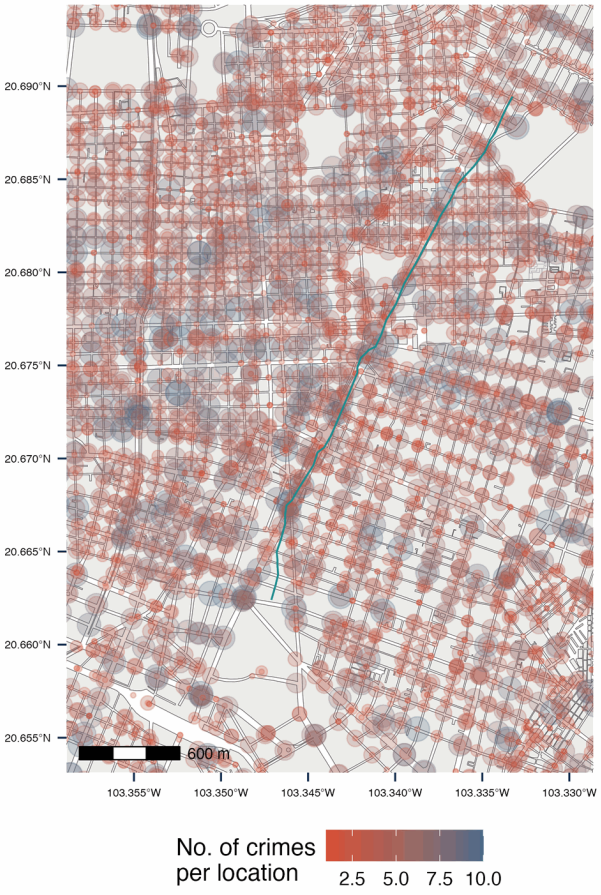


Notes: Spatial RDD estimates of the boundary effect on mean ENLACE scores in Spanish and Math at secondary and high school within 1.5 km of the boundary (2006–2014), following Calonico et al. (2014). Confidence bands are 95%. These scores begin in 2006, after the convergence visible in Figure B.5, and the estimates fluctuate around zero throughout.

B.8 Multi-wave evolution: additional outcomes

This subsection reports the outcomes omitted from the main-text multi-wave figure (Figure 4). Figure B.10 traces adult literacy across the boundary from 1930 to 2020: literacy is near-

Figure B.8: Spatial distribution of reported crimes in central Guadalajara, 2017–2024



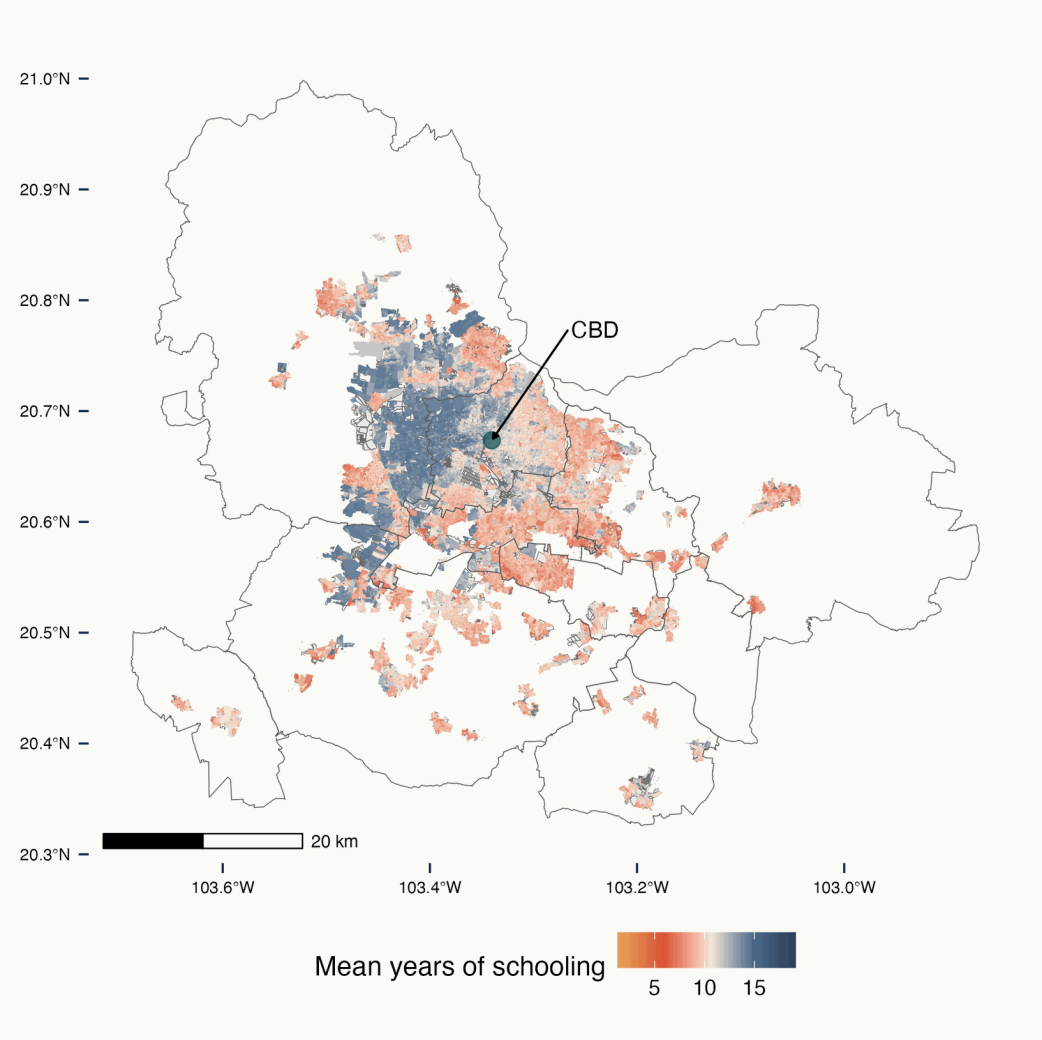
Notes: Each circle represents the count of reported crimes (*carpetas de investigación*) at a block-level coordinate, darker shades indicating higher incidence. Data include all crime categories reported to the Jalisco Attorney General's Office (2017–2024), georeferenced by IIEG. The thick green line marks the boundary. Crime is densely clustered on both sides with no visible discontinuity, confirmed formally in Table B.4.

universal on both sides throughout, and no boundary gap emerges in any wave — consistent with the human-capital divide operating through schooling and attendance rather than basic literacy.

B.9 Historical human capital gaps in 1930

The 1930 evidence highlighted in Figure 4 draws on the broader set of variables in Table B.6, restricting the sample to *cuarteles* adjacent to the boundary — cuarteles 1, 2, and 7 on the west (central Guadalajara and the historical Mexicaltzingo area) and cuarteles 8 and 9 on the east (Analco). Comparisons control flexibly for sex and age, with standard errors clustered at the household level and FDR adjustment. The pattern is consistent with the 2020 RDD. Human-capital outcomes divide the two banks sharply in 1930, with illiteracy 11.4 percentage points higher and school attendance 10.6 points lower on the eastern side, and so do occupational composition and birthplace, with Analco residents more likely to be Jalisco-born. Property ownership shows no significant differential, but the two ownership indicators are coarse and the broader one is measured on a seventh of the sample, so that null does not establish that wealth in property was equal across the banks.

Figure B.9: Mean years of schooling by census block, Guadalajara metropolitan area (2020)



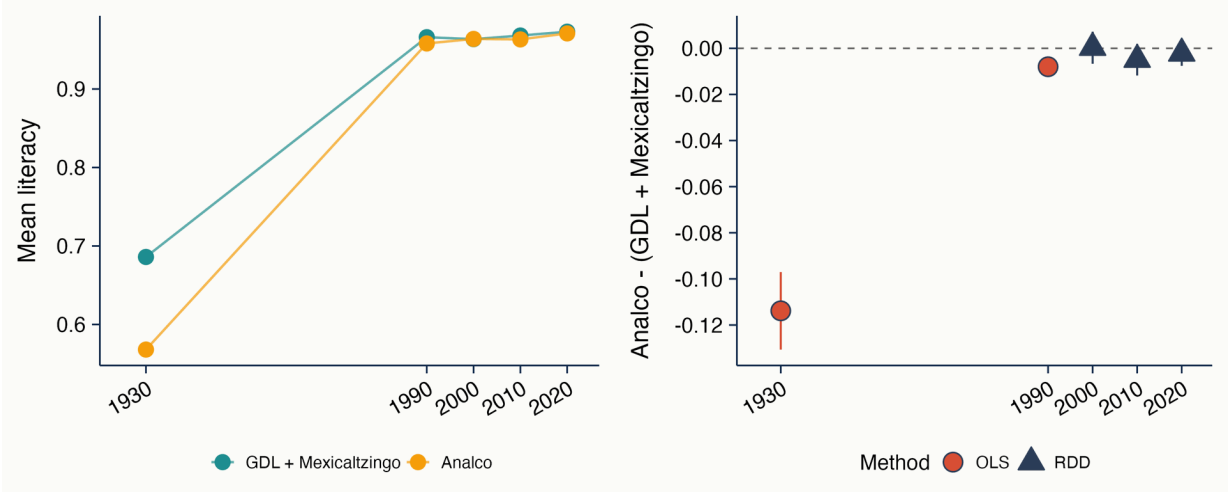
Notes: Mean years of schooling at the census-block level for the ten municipalities of the Guadalajara metropolitan area, from the public block-level statistics of the 2020 *Censo de Población y Vivienda* (INEGI). Shading runs from red (lower mean schooling) to blue (higher mean schooling), on the same scale as Figure 1b; unshaded polygons are blocks with suppressed or zero values, and dark lines are municipal boundaries. The point marks the central business district. The east–west gradient visible at the historical boundary continues across municipal lines.

B.10 Donut estimates

B.11 Placebo boundaries and recentered specification

I randomly shift and rotate the boundary 1,000 times within the municipality and re-estimate the boundary-adaptive RDD at each placebo location. Three examples are shown in Figure B.11; each preserves the length and roughly the orientation of the original but is otherwise drawn at random. The draws are generated as follows. Each placebo takes the original boundary and applies a rigid translation and a rotation about its centroid, leaving its length and shape unchanged. The longitude shift is drawn uniformly from $\pm[0.01, 0.05]$ degrees and the latitude shift from $\pm[0.01, 0.035]$ degrees, with the two signs equally likely and the interval around zero excluded, so that no placebo lies within roughly one kilometre of the true line; at this latitude the shifts span

Figure B.10: Adult literacy across the historical boundary, 1930–2020



Notes: As in Figure 4, for adult literacy. Panel (a) plots the mean literacy rate on each side of the boundary; panel (b) the Analco-minus-west gap, with OLS estimates for 1930 and 1990 (circles) and non-parametric spatial RDD estimates for 2000–2020 (triangles). In 1930 literacy stood at 68.6 percent on the western side against 56.8 on the eastern, a gap of 11.4 percentage points that is significantly different from zero. The series has no wave between 1930 and 1990, and by 1990 literacy is near-universal on both banks, close to 96 percent, with a gap under one percentage point. It remains indistinguishable from zero in 2000, 2010 and 2020.

1.0 to 5.2 km east or west and 1.1 to 3.9 km north or south. The rotation is drawn uniformly from $\{0°, 45°, 90°, 135°\}$. A draw is admissible if the shifted boundary still intersects the municipality of Guadalajara; 984 of the 1,000 draws satisfy that rule and 16 are dropped, all of them at 45° or 90°, and a draw is dropped as well if the estimator fails to return at that location. Sides are assigned by closing the shifted line into a polygon and labelling blocks by whether their centroid falls inside it, with the closing orientation set by the rotation, so that the side coding rotates with the boundary rather than being fixed to east and west. As at the true boundary, the estimator is evaluated at five regularly spaced points along the shifted line and the placebo statistic is the equally weighted average of the five pointwise effects, which keeps the support and the number of evaluation points the same across draws. Recentering follows Borusyak and Hull (2023). For each census block I compute its expected exposure across the admissible draws, meaning the share of placebo boundaries that place it on the treated side and its average signed distance to them, regress block schooling on those two variables and their interaction, and re-estimate the boundary-adaptive RDD on the residuals. The exposure measures are therefore block-level and are built from the same draws that generate the placebo distribution. Figure B.12 reports the distribution of placebo coefficients: approximately symmetric and centered near zero, with the true-boundary estimate in the left tail. Following Borusyak and Hull (2023), a recentered specification that adjusts each block's outcome for its average exposure to placebo boundaries gives a true-boundary estimate of −0.557 SD — *larger* than the unadjusted −0.499. The two-sided randomization-inference p-value is 0.094 and the left-sided p-value 0.051.

Table B.6: Socio-demographic differences across the historical boundary, 1930 Census

Variable	West Mean	East Mean	Diff.	Std. Error	q-value	N
Property: Some	0.233	0.212	-0.013	0.023	0.349	1,315
Property: Residence	0.023	0.020	-0.003	0.003	0.251	9,238
Birthplace: Jalisco	0.877	0.931	0.052	0.006	0.001	9,169
Occupation: Zapatero	0.009	0.025	0.015	0.003	0.001	9,172
Occupation: Merchant	0.050	0.052	0.002	0.004	0.390	9,172
Occupation: Laborer	0.032	0.035	0.002	0.004	0.349	9,172
Occupation: Mason	0.013	0.015	0.002	0.002	0.274	9,172
Occupation: Employed in firm	0.019	0.014	-0.005	0.003	0.082	9,172
Occupation: Maid	0.016	0.005	-0.011	0.002	0.001	9,172
Occupation: Housework	0.305	0.332	0.033	0.007	0.001	9,172
Human Capital: Illiterate	0.314	0.432	0.114	0.009	0.001	9,238
Human Capital: Studying	0.375	0.266	-0.106	0.012	0.001	3,965

Notes: OLS estimates of the east-minus-west difference in mean characteristics among residents of the *cuarteles* adjacent to the historical boundary in the 1930 Population Census of Guadalajara (Anderson and Spike, 2007). The west-bank sample includes cuarteles 1, 2, and 7 (central Guadalajara and the historical Mexicaltzingo area); the east-bank sample includes cuarteles 8 and 9 (Analco). Variables capture property ownership, birthplace (born in Jalisco), principal occupational categories, and two human-capital outcomes: adult illiteracy and current school attendance among individuals aged 18 or younger. All estimates control for sex and flexible controls for age, with standard errors clustered at the household level. The q-value column reports sharpened FDR-adjusted p-values (Anderson, 2008) across all twelve variables. The east–west gaps in birthplace, occupational structure, illiteracy, and school attendance are large and significant after correction, while the property comparisons are not. The two property variables are indicators of whether a person reports owning any property or owning their residence, not values or any wider measure of wealth, and the first is recorded for 1,315 individuals against roughly 9,200 for occupation and literacy, so the null there is imprecise and silent about property values. The 1930 census records no ethnicity. The table therefore describes where the measured differences lie at this date rather than identifying the channels through which the boundary operated.

B.12 Residential mobility and selective in-migration

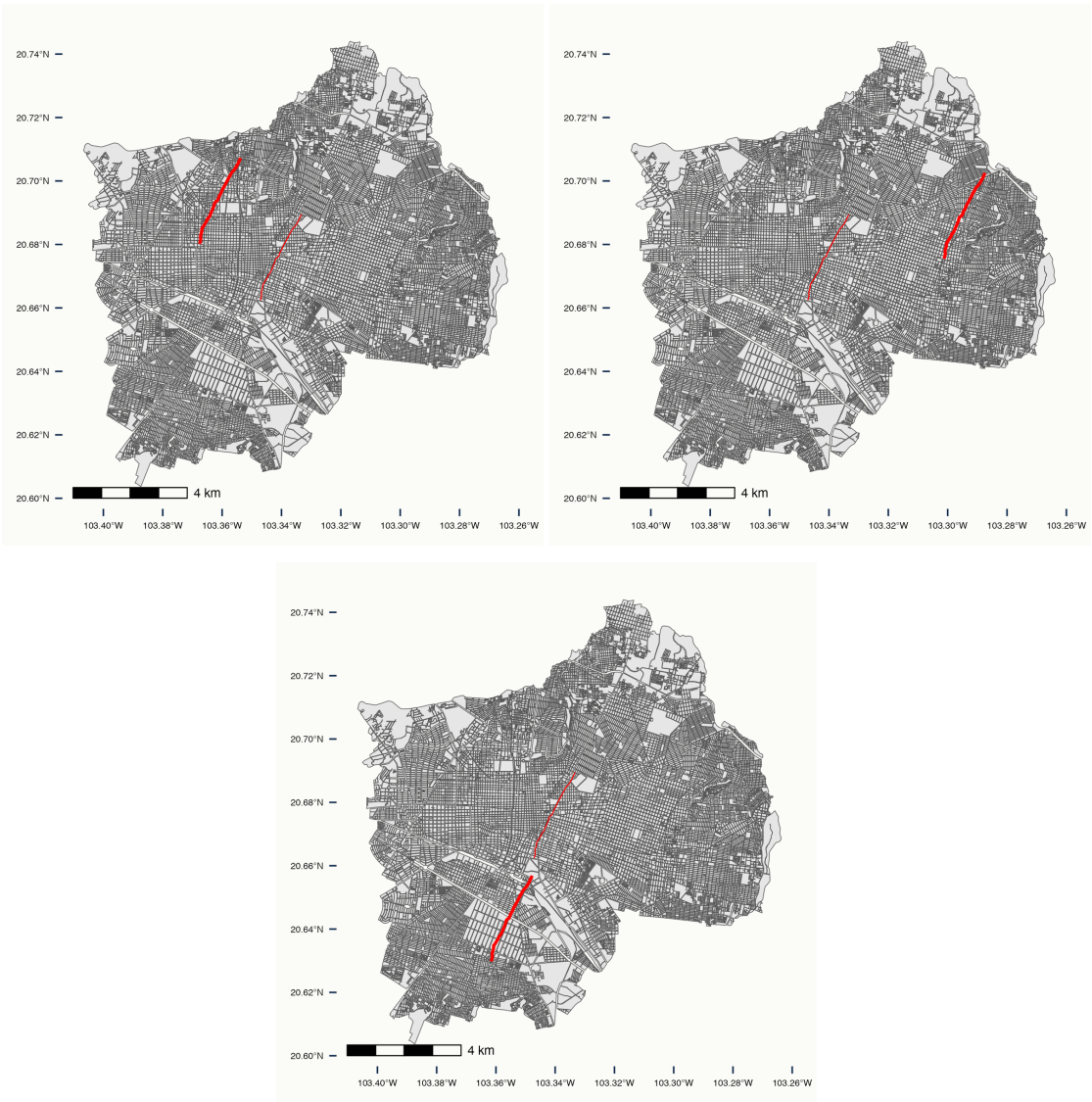
According to the 2020 census, 91.5% of Guadalajara residents lived in Jalisco five years earlier, and re-estimating the RDD excluding inter-state migrants leaves the results essentially unchanged (last row of Panel A, Table B.1). To approximate within-city mobility, I match adults aged 34+ in 2020 to plausibly the same individuals in the same block in 2010, using nearest-neighbor distance on age, sex, years of schooling, and birthplace, and classify an individual as a long-term resident if the match distance falls below the median. Table B.8 reports the boundary effect on that indicator, which is positive and significant once the clean controls are added. Two features of the construction limit what the indicator can carry. The census records are anonymous, so the match links a 2020 resident to a demographically similar 2010 resident of the same block rather than to the same person, and a block whose households turned over completely would still return

Table B.7: Donut RDD estimates on human capital, 2020 Census

$\hat{\tau}$	SE	95% CI	p-val	h	N_h	Donut (m)
Panel A. Years of schooling (ages 6+)						
-0.165	0.033	[-0.229, -0.101]	0.000	570.8	25011	0
-0.101	0.026	[-0.151, -0.050]	0.000	976.0	57387	50
-0.211	0.031	[-0.272, -0.149]	0.000	667.2	32464	100
-0.151	0.031	[-0.212, -0.091]	0.000	916.7	49857	150
Panel B. School attendance (ages 6–22)						
-0.075	0.045	[-0.163, 0.013]	0.094	1200.2	18265	0
-0.086	0.047	[-0.178, 0.007]	0.069	1022.6	14606	50
-0.100	0.048	[-0.194, -0.006]	0.037	1098.3	15686	100
-0.127	0.056	[-0.237, -0.016]	0.025	954.9	12429	150

Notes: Spatial RDD estimates excluding all residents within the stated distance of the historical boundary, in standardized units, following Calonico et al. (2014) with a uniform kernel, local linear polynomial and the NN variance estimator. Every row, including the zero-donut row, is estimated without covariates, since covariate coverage is thinnest in the blocks nearest the boundary; the zero-donut rows are the no-covariate specifications of Table B.1. The bandwidth is re-selected for each radius and moves substantially across rows, so the variation in the estimates reflects both the exclusion and the width of the estimation window. The table is therefore informative about the stability of the sign rather than a controlled comparison isolating the innermost observations. Residential density at the cutoff is far lower on the western bank, which is Guadalajara's historic commercial core, so the innermost observations come disproportionately from one side. The two densities converge only at about 500 meters (Figure B.1), so all three radii fall inside the range over which they differ. The donut removes the observations closest to the line without reaching the distance at which the two banks are comparably populated.

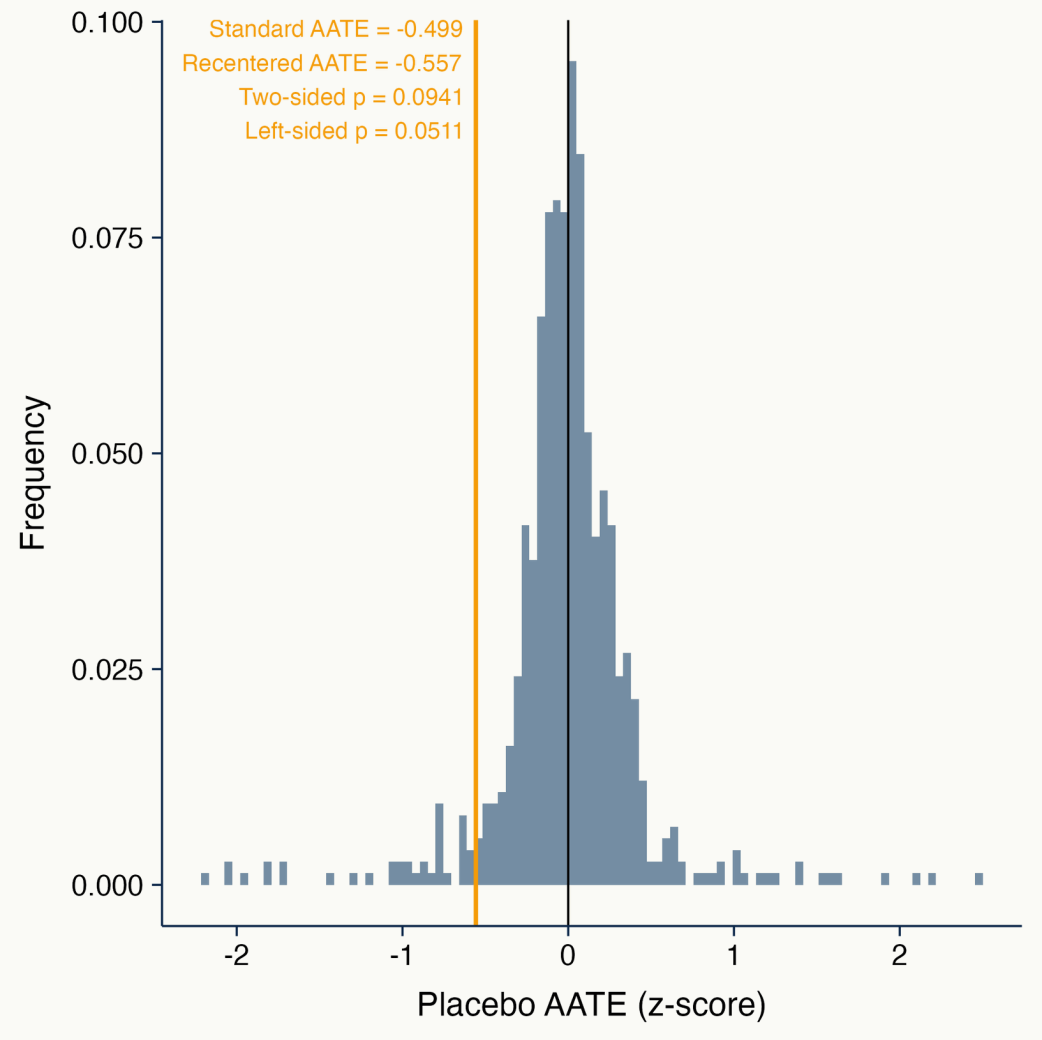
Figure B.11: Examples of placebo boundaries for the spatial falsification test



Notes: Three illustrative examples (out of 1,000) of placebo boundaries used in the spatial falsification test of Figure B.12. Each panel shows one placebo realization (thick red line) generated by randomly shifting and rotating the historical SJdD creek boundary (thin red line) within Guadalajara. Placebos preserve the length and approximate orientation of the original but are otherwise random. Each placebo is used to estimate a boundary-adaptive RDD on census-block mean years of schooling; the resulting distribution is in Figure B.12.

close matches if the arrivals resembled the departures. The nearest-neighbour distance also falls mechanically with the number of 2010 records available in the block, and residential density at the boundary is higher on the eastern side (Figure B.1), so eastern blocks offer larger candidate pools and shorter distances for that reason alone. That bias runs in the same direction as the estimate. I therefore do not read the coefficient as a measured difference in residential turnover, and the argument below rests on the migrant-exclusion row rather than on it. What the two together weigh against is one specific account of the gap, in which the eastern side looks worse because it has absorbed recent and less-educated arrivals. Dropping inter-state migrants moves the estimates hardly at all, and nothing in the mobility evidence points to the excluded side being the more transient one. The model does not derive that pattern either. Churn enters it as an exogenous parameter whose effect on local composition the model traces, and not the reverse. The remark

Figure B.12: Distribution of placebo-boundary treatment effects with recentering



Notes: Distribution of standardized boundary-adaptive RDD estimates from 1,000 placebo boundaries (Figure B.11), applied to census-block mean years of schooling in the 2020 census for Guadalajara. The horizontal axis is the estimated effect in standardized units. The black vertical line at zero marks the null. The unadjusted true-boundary estimate is −0.499 SD and lies in the left tail. Following Borusyak and Hull (2023), the orange line marks the recentered estimate (−0.557 SD), which adjusts each block's outcome for its average exposure to the placebo boundaries, removing geographic-exposure bias. The two-sided randomization-inference p-value is 0.094 and the left-sided p-value 0.051. The exercise rules out that boundary-magnitude discontinuities of this size are common at arbitrary locations.

in Appendix D.9 notes that type-1 agents on the low branch have an additional reason to stay, a channel that would run in this direction if residential choice were modeled explicitly, but it does not identify rootedness as the source of persistence. It does not rule out selective migration more generally. What matters for the schooling gap is the educational composition of those who arrive and those who leave, not the amount of turnover, so a small but strongly selected flow could still open a compositional difference. The census records nothing about former residents, so the selection of leavers is unobserved here and this exercise bounds neither margin. The measure also refers to one period. It compares residents in 2020 with residents of the same block in 2010, so it describes turnover over the last decade. Guadalajara's population was reshaped earlier, by the rural-to-urban migration of the twentieth century, and that episode is the one the model's race between local transmission and turnover concerns. Neither this measure nor the 1821 migrant

shares of Appendix A.6 observes it, and the two sit on either side of it.

Table B.8: Spatial RDD estimates on probability of long-term residence (individual level, 2020 Census)

Outcome	$\hat{\tau}$	SE	95% CI	p-val	Poly	h	b	N_h	N_b
Probability to stay in place	0.144	0.075	[-0.003, 0.291]	0.054	1	131.9	254.9	1,315	3,672
Probability to stay in place (with controls)	0.178	0.080	[0.022, 0.334]	0.026	1	116.3	208.9	1,121	2,605

Notes: Spatial RDD estimates of the boundary effect on the probability of long-term residence, individual-level 2020 census data for Guadalajara, restricted to adults aged 34+. The outcome is constructed via nearest-neighbor matching of each 2020 resident to a 2010 resident of the same block on age, sex, years of schooling, and birthplace; an individual is a long-term resident if the match distance is below the median. Inference follows Calonico et al. (2014); the control specification adds sex, age, longitude, latitude, and boundary-segment fixed effects. The effect is positive and significant once the controls are added: residents east of the boundary are about 18 percentage points more likely to be long-term residents. This works against a selective-mobility explanation — the lower-attainment side is the more rooted one.

B.13 The 1992 explosions: design details

This subsection describes the design behind Figure 10 in Section 5.5. A block is directly affected if it intersects an 80-meter buffer around the collector's path, digitized from a georeferenced photomosaic of the damaged corridor. First-degree neighbors of the affected blocks are excluded from the analysis as a buffer zone; second- through fifth-degree neighbors form the comparison group. The sample is restricted to the east bank, within two kilometers of the historical boundary. The 2000–2020 estimates use individual-level census microdata at the block level, accessed through INEGI's microdata laboratory. Each wave-specific coefficient compares residents of affected and comparison blocks, with age and sex fixed effects and standard errors clustered by block; the estimates are unweighted. Years of schooling are measured among adults aged 18 and older. School attendance is measured among children aged 6–18, conditional on parental years of schooling. The 1990 census is geolocated at the AGEB level, so the pre-shock placebo compares the AGEBs intersecting the affected corridor with the remaining east-bank AGEBs under the same specification, clustering by AGEB.

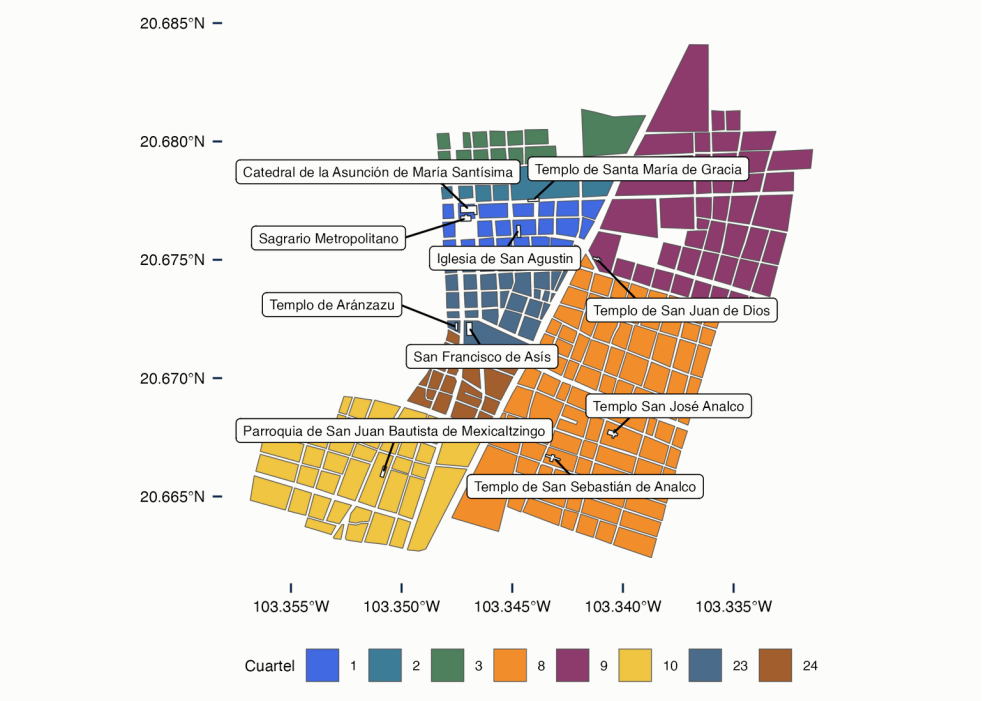
C Neighborhood Social Differences: Supporting Material

This appendix collects the supporting material for Section 4, in the order the section draws on it. Appendix C.1 maps the parishes behind the patron-saint assignment and Appendix C.2 reports the calendar-saint placebo. Appendix C.3 reports the additional Life Quality Survey results, and Appendix C.4 the moderated mediation analysis that earlier drafts carried in the main text.

C.1 Parishes and patron saints around the creek

Figure C.1 maps the *cuarteles* adjacent to the SJdD creek and the temples of their parishes. The patron-saint assignment used in Section 4.2 follows this map. Cuarteles 8 and 9 form Analco on the eastern bank, cuartel 10 is Mexicaltzingo, and the remaining cuarteles lie on the western bank.

Figure C.1: *Cuarteles* and parish temples around the San Juan de Dios creek



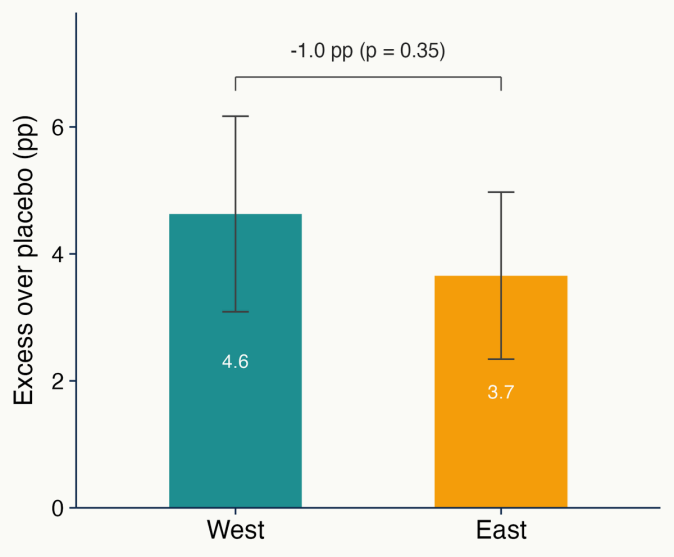
Notes: *Cuarteles* adjacent to the SJdD creek in the cartography of the Guadalajara Census Project (Anderson and Spike, 2007), colored by cuartel number, with the parish temples overlaid and labeled (building footprints from OpenStreetMap). Each cuartel is assigned the patron saints of its parish temples; the patron-saint exercise in Section 4.2 uses this assignment.

C.2 Calendar-saint naming: the santoral placebo

Figure C.2 reports the calendar-saint exercise referenced in Section 4.2. In the baptismal registers, naming a child after the saint of the baptism day is a common practice on both banks, and equally so: the excess over the shifted-calendar placebo is 3.7 percentage points on the east and 4.6

on the west, a difference not distinguishable from zero ($p = 0.35$). The placebo indicates that general devotional naming was similar on both banks, so the patron-saint difference documented in Section 4.2 is specific to the local parish.

Figure C.2: Naming after the saint of the baptism day



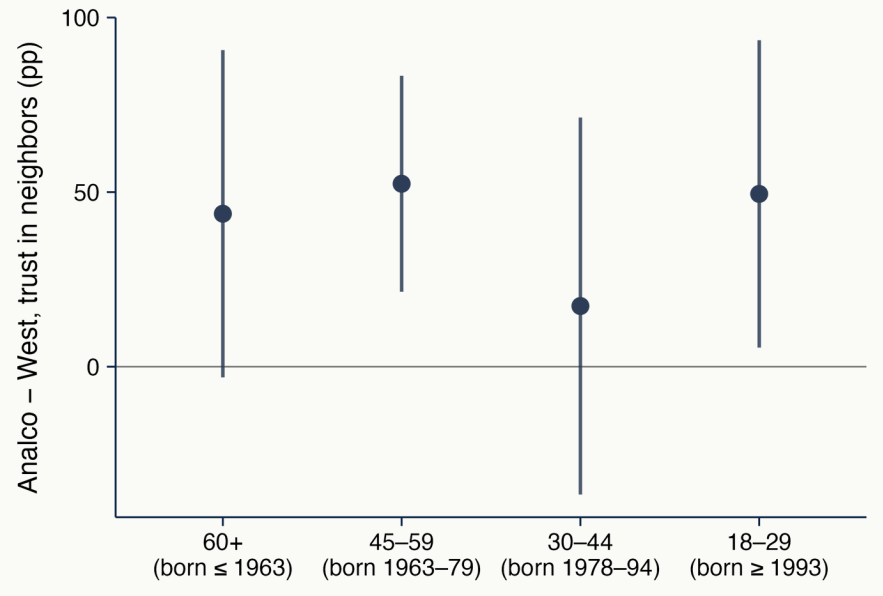
Notes: FamilySearch baptismal registers of the parishes adjacent to the creek, 1850–99 (east: Analco and San Juan de Dios; west: Sagrario, El Pilar, and Guadalupe). Each bar reports the probability that a child's given name matches a saint of the baptism day, in excess of the mean match under seven placebo calendars shifted by 45 to 321 days, with 95 percent confidence intervals. The comparison of name and date is within record, and the placebo absorbs record-level transcription conventions, including the number of name tokens transcribed, so the excess is comparable across parishes. The santoral covers 105 feast days and 170 name keys; 23 percent of baptisms fall on covered days ($N = 3{,}298$). Bar labels report the excess in percentage points, and the bracket reports the east–west difference with its p-value.

C.3 Additional Life Quality Survey results

C.3.1 The neighbor-trust gap by age bracket

Figure C.3 estimates the Analco gap in trust in neighbors separately by the Life Quality Survey's four age brackets, with the same controls as the main specification. The gap is present in every bracket, at about 50 percentage points among respondents aged 18–29, born after the early 1990s, and it is statistically indistinguishable across the four. The survey records age only in these brackets, and with two waves age and birth cohort cannot be separated, so the figure documents a gap of similar size at every age observed rather than a cohort trend. The youngest bracket matters for the argument because it shows the gap alive among respondents schooled long after the provision differences of Section 4.1 had closed.

Figure C.3: The Analco gap in trust in neighbors, by age bracket



Notes: Each point reports the Analco coefficient from the trust-in-neighbors specification of Table 2, estimated separately within each age bracket of the Life Quality Survey (sample sizes 122–204), with sex, survey-year, and years-of-schooling fixed effects and side-specific distance gradients; vertical bars show 95 percent heteroskedasticity-robust confidence intervals. Approximate birth cohorts follow from the 2022 and 2024 survey years. The four brackets are estimated on disjoint subsamples, so the gaps are independent. A Wald test does not reject equality across them ($W = 1.27$ on three degrees of freedom, $p = 0.74$), while the hypothesis that all four equal zero is rejected ($W = 19.6$ on four degrees of freedom, $p < 0.001$). With a cross-section of two waves, age and cohort effects cannot be separated.

C.3.2 Institutional trust

Table C.1 reports the institutional-trust index (Col. 1) and disaggregates it into its ten items (Cols. 2–11). Residents east of the boundary report lower trust in every institution: nine of the ten items are significant at the 5 percent level and the prosecutor's office at 10 percent. The largest gaps are for the Navy (−47.3 pp) and the Army (−39.7 pp).

C.3.3 Social participation

Table C.2 disaggregates participation. Residents on the eastern side are no more likely to belong to any organization (Col. 1) and report similar numbers of memberships (Col. 2), but are 8.5 pp *more* likely to participate in a neighborhood or residents' association (Col. 3) — about twice the sample mean — with no significant gap for school-related, political, or religious membership, and if anything lower NGO membership (Cols. 4–7). The asymmetry is the participation analog of the trust pattern: civic engagement is dense but bounded to the immediate neighborhood.

Table C.1: Localized social capital: trust in institutions

	Institutional trust index	Trust in churches	Trust in local congress	Trust in electoral institute	Trust in political parties	Trust in government	Trust in judiciary	Trust in police	Trust in army	Trust in navy	Trust in prosecutor's office
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\mathbb{1}\{Analco\}$	-0.247***	-0.205**	-0.204***	-0.257***	-0.168***	-0.258***	-0.217***	-0.263***	-0.397***	-0.473***	-0.113*
	(0.046)	(0.084)	(0.060)	(0.072)	(0.046)	(0.054)	(0.054)	(0.060)	(0.079)	(0.080)	(0.059)
Observations	578	578	578	578	578	578	578	578	578	578	578
Dep. var. mean	0.2297	0.4810	0.1652	0.2647	0.1436	0.1773	0.1990	0.2007	0.3469	0.3728	0.2042
R^2	0.1102	0.1679	0.0751	0.0952	0.0902	0.0870	0.0970	0.0763	0.1493	0.1651	0.0911
q-value	[0.001]	[0.019]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.045]
Age-Sex fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Years of schooling fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: OLS estimates from the 2022 and 2024 waves of the local Life Quality Survey (*Jalisco Cómo Vamos*) for the municipality of Guadalajara. Each column is a separate 0–1 institutional-confidence outcome; column (1) is the institutional-trust index (first principal component of the ten items). All specifications include age-by-sex and survey-year fixed effects, absorb years of schooling, and control for distance to the boundary. Heteroskedasticity-robust standard errors in parentheses. Sharpened FDR q-values (Anderson, 2008), adjusted across all social-capital outcomes, in brackets.

Table C.2: Localized social capital: organizational membership

	Member of any organization	Number of memberships	Member of neighborhood assoc.	Member of school-related org.	Member of political org.	Member of religious group	Member of NGO/ assistance org.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\{Analco\}$	-0.078	-0.040	0.085**	-0.020	0.035	-0.094	-0.047
	(0.074)	(0.139)	(0.041)	(0.032)	(0.033)	(0.068)	(0.045)
Observations	578	578	578	578	578	578	578
Dep. var. mean	0.1176	0.1834	0.0415	0.0173	0.0381	0.0692	0.0173
R^2	0.1589	0.1219	0.0879	0.0641	0.1225	0.1163	0.0647
q-value	[0.111]	[0.272]	[0.035]	[0.188]	[0.111]	[0.087]	[0.111]
Age-Sex fixed effects	✓	✓	✓	✓	✓	✓	✓
Survey year fixed effects	✓	✓	✓	✓	✓	✓	✓
Years of schooling fixed effects	✓	✓	✓	✓	✓	✓	✓

Notes: OLS estimates from the 2022 and 2024 waves of the local Life Quality Survey (*Jalisco Cómo Vamos*) for the municipality of Guadalajara. Columns (1)–(2) report any-membership and the count of memberships; columns (3)–(7) report membership in specific organization types. All specifications include age-by-sex and survey-year fixed effects, absorb years of schooling, and control for distance to the boundary. Heteroskedasticity-robust standard errors in parentheses. Sharpened FDR q-values (Anderson, 2008), adjusted across all social-capital outcomes, in brackets.

C.4 Neighbor trust and the educational gap: heterogeneity and falsification

This subsection reports estimates that, in earlier drafts, appeared in the main text as a moderated mediation analysis. The mediation decomposition rested on a mediator that is neither randomized nor plausibly exogenous, so I no longer report it. What the estimates can do is describe where the educational gap is larger and rule out two accounts of it. Table C.3 reports them; the causal evidence on salience is the experiment of Section 5.2.

Table C.3: Neighbor trust and educational attainment: heterogeneity and falsification

	Years of schooling	Trust in neighbors	Neighbor vs. nonlocals trust	Years of schooling					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathbb{1}\{Analco\}$	-1.95**	0.370***	0.092**	-0.864	0.079	-1.39	-2.24***	-0.859	-1.88**
	(0.795)	(0.104)	(0.043)	(0.861)	(1.21)	(0.899)	(0.811)	(0.755)	(0.785)
Years of schooling		-0.002	-0.003						
		(0.005)	(0.002)						
Trust in neighbors				0.714					
				(0.473)					
Trust in neighbors × Analco				-1.68**					
				(0.662)					
Neighbor vs. nonlocals trust					0.620				
					(1.14)				
Neighbor vs. nonlocals trust × Analco					-3.08**				
					(1.52)				
Institutional trust index						1.62			
						(1.16)			
Institutional trust index × Analco						-1.73			
						(1.52)			
Tract-agg. neighbor trust (z)							0.712***		
							(0.216)		
Tract-agg. neighbor trust (z) × Analco							-1.41***		
							(0.328)		
Tract-agg. schooling (z)								1.38***	
								(0.195)	
Tract-agg. schooling (z) × Analco								0.151	
								(0.406)	
Tract-agg. land value (log, z)									1.09***
									(0.187)
Tract-agg. land value (log, z) × Analco									-0.680
									(0.499)
Observations	578	578	578	578	578	578	578	578	578
Dep. var. mean	11.34	0.6678	0.6003	11.34	11.34	11.34	11.34	11.34	11.34
R^2	0.2124	0.0866	0.0824	0.2207	0.2209	0.2153	0.2352	0.2999	0.2545
Age-Sex fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimates from the 2022 and 2024 waves of the local Life Quality Survey (*Jalisco Cómo Vamos*) for Guadalajara. All models include age-by-sex and survey-year fixed effects and control for distance to the boundary and its interaction with the Analco indicator. The dependent variable is years of schooling in columns (1) and (4)–(9); columns (2)–(3) report the Analco gap in the two trust measures themselves. Columns (4)–(6) interact each trust measure with the Analco indicator, so the main effect is the western slope and the interaction the difference between the two sides. Columns (7)–(9) interact the Analco indicator with tract averages of neighbor trust, completed schooling, and log land values. Heteroskedasticity-robust standard errors in parentheses. The trust measures are not randomized, so these estimates describe heterogeneity and bound two accounts of the gap; they do not identify a channel.

Individual mediation. Columns (4)–(6) interact the Analco indicator with candidate mediators, so the reported main effect is the western slope and the interaction is the difference between the two sides rather than the eastern slope itself. In Column (4) neighbor trust is modestly positively associated with schooling on the western side, at 0.714, and the interaction of −1.68 implies an eastern slope of about −0.97: among Analco residents, higher trust in neighbors goes with

roughly one fewer year of schooling. The neighbor-vs-nonlocals premium in Column (5) gives a larger difference, −3.08 against a western slope of 0.620, for an eastern slope near −2.46. The institutional-trust index in Column (6) gives an interaction of −1.73 that is not statistically distinguishable from zero, and a western slope of 1.62, so the implied eastern slope is close to zero. The indirect effects reported below are computed from the side-specific slopes rather than from the interactions.

Moderated mediation. Columns (2) and (3) report the Analco gap in the two trust measures themselves, which is the association documented in Section 4.3: residents of the excluded side report higher trust in neighbors and a higher neighbor-versus-nonlocals premium.

Contextual moderators. Columns (7) to (9) interact the Analco indicator with three tract-level variables. The gap is larger where average neighbor trust in the tract is higher (interaction −1.41), so the pattern within individuals also appears between neighborhoods. The remaining two columns rule out accounts of the gap rather than supporting one. Interacting with average completed schooling in the tract gives a null (interaction 0.15), so the gap is not a matter of individuals converging on the schooling their neighbors already have. Interacting with log land values also gives a null (interaction −0.68), so it is not wealth-based sorting across the boundary either. These are tract averages of realized outcomes rather than measurements of the state variables of Appendix D, and the nulls bound the two accounts weakly given the size of the Life Quality Survey. Tract-level confounders correlated with both neighbor trust and schooling could produce the first pattern.

D Theoretical Framework

This appendix develops the framework used in Section 5.1. The model separates public educational provision P_ℓ, educational investment e_i, and a neighborhood state q_ℓ summarizing the share of residents for whom local social relations enter educational choices.

Historical exclusion enters through unequal provision. Poorer provision raises the effective cost of educational investment and can generate an initial schooling difference. Once neighborhoods are on different education paths, the incentives to transmit neighborhood orientation can differ across them. The resulting change in q_ℓ feeds back into educational choices. After the provision-related cost wedge disappears, this local state can remain different when transmission is strong relative to residential turnover.

The baseline does not assume that public provision changes the value of local networks. I consider that possibility separately in Appendix D.10.

The state q_ℓ is a modeling device, not an empirical latent index. The historical naming records and present-day surveys in Section 4 measure different expressions of local social orientation in different periods. I do not treat them as repeated measures of the same q_ℓ, nor do I estimate q_ℓ from them.

D.1 Setup and educational choice

Population. Consider two neighborhoods, A and B, in the same city. Agents have type

$$\theta_i \in \{0, 1\}.$$

Type-1 agents place weight on local social relations and on conformity with local educational behavior. Type-0 agents do not. The type is a device for representing neighborhood-oriented behavior; it is not an ethnic marker or a direct empirical measure of neighborhood orientation.

Let q_ℓ be the type-1 share in neighborhood ℓ, and let $\bar{e}_\ell$ denote average educational investment. The two neighborhoods start with the same composition:

$$q_A^0 = q_B^0 = \bar{q}, \tag{D.1}$$

where $\bar{q} \in (0, 1)$. The model therefore does not require an ex-ante difference in neighborhood orientation.

Provision and preferences. Public educational provision in neighborhood ℓ is P_ℓ. I interpret P_ℓ narrowly as provision affecting the cost and feasibility of schooling: access, availability, travel, direct costs, and related frictions.

An agent chooses $e_i \in \{0, 1\}$, where $e_i = 1$ denotes educational investment. Utility is

$$U_i(e_i \mid \theta_i, q_\ell, \bar{e}_\ell) = w_0 + \big[\rho - C(P_\ell)\big]e_i + a(s)\theta_i \left[\Gamma q_\ell(1 - \delta e_i) - \kappa|e_i - \bar{e}_\ell|\right]. \tag{D.2}$$

The parameter $\rho > 0$ is the private return to education and is common across neighborhoods. Provision affects the cost term:

$$C'(P) < 0. \tag{D.3}$$

Better provision lowers the effective cost of educational investment. The baseline does not assume that P_ℓ changes ρ.

For notation, define

$$r_\ell \equiv \rho - C(P_\ell). \tag{D.4}$$

The remaining parameters are $\Gamma > 0$, the value of local-network interaction; $\delta \in (0, 1)$, the share of this payoff crowded out by education; and $\kappa > 0$, the cost of departing from local educational behavior.[28] The term $a(s) > 0$ gives the weight placed on the social component when the neighborhood has salience s. I normalize $a = 1$ outside the salience exercise.

In the baseline,

$$C'(P) < 0, \qquad \frac{\partial \rho}{\partial P} = 0, \qquad \frac{\partial \Gamma}{\partial P} = 0.$$

A type-0 agent invests whenever $r_\ell \geq 0$. For a type-1 agent,

$$U_i(1) - U_i(0) = r_\ell - a(s)\left[\delta\Gamma q_\ell + \kappa(1 - 2\bar{e}_\ell)\right]. \tag{D.5}$$

Define the *neighborhood pull*

$$\Pi_\ell \equiv \delta\Gamma q_\ell + \kappa(1 - 2\bar{e}_\ell). \tag{D.6}$$

A type-1 agent invests if and only if

$$\rho - C(P_\ell) \geq a(s)\Pi_\ell. \tag{D.7}$$

Provision and salience act on different margins. Better provision lowers the effective cost on the left-hand side. Salience changes the weight placed on local considerations on the right-hand side. The network component of Π_ℓ depends on who lives in the neighborhood, while the conformity component depends on what residents do, in the terminology of Manski (1993).

D.2 Salience

Suppose the salience treatment changes attention from $a(0)$ to

$$a(1) = a(0) + \Delta a, \qquad \Delta a > 0.$$

For a type-1 agent, the treatment changes the utility difference between educational investment and non-investment by

$$\Delta\big[U_i(1) - U_i(0)\big] = -\Delta a\, \Pi_\ell. \tag{D.8}$$

[28] With binary education, the absolute-deviation term produces the same type-1 utility difference as the corresponding quadratic social-interactions specification. See (Glaeser and Scheinkman, 2001; Scheinkman, 2008). The conformity term is also related to identity-based schooling costs in Akerlof and Kranton (2002).

The sign depends on the local state. Salience lowers the relative payoff to education when $\Pi_\ell > 0$ and raises it when $\Pi_\ell < 0$. The choice is binary, so the model delivers the sign of the change in the utility difference and not the size of the aggregate response. With a threshold common to all agents a small Π_ℓ would move either everyone or no one; the share that switches scales with $\Delta a\, \Pi_\ell$ only once thresholds are dispersed across agents, and then it also depends on how many of them sit near indifference. The west-bank estimates are therefore uninformative about the size of Π_ℓ there rather than evidence that it is small. Equation (D.8) concerns the latent payoff difference. Without heterogeneity, all type-1 agents in a neighborhood make the same choice, so the model pins down only the sign of the demand response. With dispersion across type-1 agents in r_i or in the components of the pull, the treatment moves only agents near the investment margin, and the shifted mass grows with $|\Delta a\, \Pi_\ell|$; this is the sense in which a weak pull implies a small demand response.

Equation (D.8) is a model object. The experiment in Section 5.2 identifies the intention-to-treat effect of randomized neighborhood salience on educational demand. It does not separately identify Δa, Π_ℓ, θ_i, q_ℓ, Γ, or κ, and it does not distinguish the network component from the conformity component.

D.3 Educational regimes

Consider within-period equilibria at $a = 1$ and $r_\ell > 0$. I restrict attention to equilibria in which agents of the same type take the same action.

Proposition 1 (Educational regimes)**.** *Define*

$$q^*(P_\ell) \equiv \frac{\rho - C(P_\ell) + \kappa}{\delta\Gamma}, \qquad q^{**}(P_\ell) \equiv \frac{\rho - C(P_\ell) + \kappa}{\delta\Gamma + 2\kappa}. \tag{D.9}$$

Then $q^{**}(P_\ell) < q^*(P_\ell)$, and:

1. *a high-education equilibrium with $\bar{e}_\ell = 1$ exists if and only if*

$$q_\ell \le q^*(P_\ell);$$

2. *a low-education equilibrium with*

$$\bar{e}_\ell = 1 - q_\ell$$

exists if and only if

$$q_\ell \ge q^{**}(P_\ell);$$

3. *both equilibria exist for*

$$q_\ell \in [q^{**}(P_\ell), q^*(P_\ell)].$$

Both thresholds rise with provision:

$$\frac{dq^*}{dP} = \frac{-C'(P)}{\delta\Gamma} > 0, \qquad \frac{dq^{**}}{dP} = \frac{-C'(P)}{\delta\Gamma + 2\kappa} > 0. \tag{D.10}$$

Proof. In the high-education equilibrium, $\bar{e}_\ell = 1$ and

$$\Pi_\ell = \delta\Gamma q_\ell - \kappa.$$

A type-1 agent invests if

$$r_\ell \geq \delta\Gamma q_\ell - \kappa,$$

which is equivalent to $q_\ell \leq q^*(P_\ell)$. Type-0 agents also invest because $r_\ell > 0$.

In the low-education equilibrium, type-0 agents invest and type-1 agents do not, so

$$\bar{e}_\ell = 1 - q_\ell.$$

Then

$$\Pi_\ell = (\delta\Gamma + 2\kappa)q_\ell - \kappa.$$

Type-1 abstention requires

$$r_\ell \leq (\delta\Gamma + 2\kappa)q_\ell - \kappa,$$

or $q_\ell \geq q^{**}(P_\ell)$. The ordering follows from the common numerator and the larger denominator in q^{**}. Differentiating (D.9) gives (D.10). □

Lower provision shifts both thresholds to the left. It therefore makes the low-education equilibrium available at lower values of q_ℓ and reduces the range over which the high-education equilibrium is available. Provision does not move Π_ℓ directly in the baseline.

On the interior of the multiplicity region there is also an equilibrium outside this class, in which type-1 agents split between the two actions and are indifferent, with

$$\bar{e}_\ell = \frac{\delta\Gamma q_\ell + \kappa - r_\ell}{2\kappa} \in (1 - q_\ell, 1).$$

It is not robust: because choices are strategic complements, a small rise in local investment lowers the pull and draws in further investment, and conversely, so myopic adjustment moves away from it. I do not use this equilibrium below.

If educational investment is literally unavailable, $e_i = 1$ is outside the feasible choice set and the only possible education state is $\bar{e}_\ell = 0$. A negative value of r_ℓ by itself is not used below to claim uniqueness of that state.

D.4 Historical provision and the initial education gap

During the exclusion era,

$$P_A < P_B,$$

so

$$C(P_A) > C(P_B), \qquad r_A < r_B.$$

The historical argument does not require the model to assume that residents of A began with a different q. The initial difference comes from schooling.

One limiting case is a period in which schooling is effectively unavailable. This generates a low initial education state without any social mechanism. The historical maps in Section 4.1 document a large asymmetry in named educational provision; they do not establish the literal absence of every educational opportunity. I therefore use infeasibility only as a limiting case.

Once educational investment is feasible but remains relatively costly in A, the regime conditions in Proposition 1 apply. Suppose

$$0 < r_A < r_B \qquad \text{and} \qquad \bar{q} < q^{**}(P_B). \tag{D.11}$$

Then only the high-education equilibrium exists in B.

In A, a sufficiently low r_A can make the low-education equilibrium unique:

$$r_A < \delta\Gamma\bar{q} - \kappa \quad \Longleftrightarrow \quad \bar{q} > q^{*}(P_A). \tag{D.12}$$

If instead

$$q^{**}(P_A) \leq \bar{q} \leq q^{*}(P_A), \tag{D.13}$$

both equilibria are available. An inherited low education norm selects the low branch under myopic best responses. Starting near $\bar{e}_A = 0$, type-0 agents invest when $r_A > 0$, while type-1 agents remain out when the low-branch condition holds. The resulting norm is

$$\bar{e}_A = 1 - \bar{q}.$$

The inherited norm can itself be generated in the model by an earlier period in which schooling was sufficiently costly or infeasible. This selection does not require an ex-ante difference in resident types.

At this point,

$$\bar{e}_A < \bar{e}_B, \qquad q_A = q_B = \bar{q}.$$

Unequal provision has generated an education difference, but not yet a difference in neighborhood orientation.

D.5 Cultural transmission

I use the cultural-transmission framework of Bisin and Verdier (2000, 2001). Each parent has one child and chooses socialization effort $\tau_\theta \in [0, 1]$ at cost

$$c(\tau) = \frac{\psi}{2}\tau^2, \qquad \psi > 0.$$

With probability τ_θ, the child adopts the parent's type. Otherwise the child draws a type from the local population. The transition probabilities are

$$\Pr(1 \mid 1) = \tau_1 + (1 - \tau_1)q, \tag{D.14}$$

$$\Pr(1 \mid 0) = (1 - \tau_0)q. \tag{D.15}$$

Parents are imperfectly empathetic: they evaluate the child's future actions using their own preferences. Let $V_r^{\theta\theta'}$ denote the value that a type-θ parent assigns to a type-θ' child in education regime r, and define

$$\Delta V_r^{\theta} \equiv V_r^{\theta\theta} - V_r^{\theta\theta'}, \qquad \theta' \neq \theta. \tag{D.16}$$

For interior effort choices,

$$\tau_1 = \frac{(1-q)\Delta V_r^1}{\psi}, \qquad \tau_0 = \frac{q\Delta V_r^0}{\psi}, \tag{D.17}$$

with the corresponding 0 and 1 corners when the unconstrained solutions lie outside $[0, 1]$.

Aggregating across parents gives

$$q_{t+1} - q_t = q_t(1-q_t)\left[\tau_1(q_t) - \tau_0(q_t)\right] \equiv \Phi_r(q_t). \tag{D.18}$$

The terms $1-q$ and q in (D.17) are the probabilities that failed vertical transmission exposes a child to the other type. This is the cultural-substitution force in Bisin and Verdier (2001). Its direction here also depends on the parental stakes, which come from the education problem.

D.6 Transmission stakes

Let

$$s_\Gamma \equiv \delta\Gamma + 2\kappa \tag{D.19}$$

and write $r = \rho - C(P)$.

In the high-education equilibrium, both types choose $e = 1$. Because the child's educational action is the same under either type, the minimal specification gives

$$\Delta V_H^1 = \Delta V_H^0 = 0. \tag{D.20}$$

The same is true when education is infeasible and both types choose $e = 0$.

In the low-education equilibrium, type-1 agents choose $e = 0$ and type-0 agents choose $e = 1$. A type-1 parent values these two outcomes as

$$\begin{aligned} V_L^{11} &= w_0 + \Gamma q - \kappa(1-q), \\ V_L^{10} &= w_0 + r + (1-\delta)\Gamma q - \kappa q. \end{aligned}$$

Hence

$$\Delta V_L^1 = s_\Gamma q - \kappa - r. \tag{D.21}$$

A type-0 parent assigns value $w_0 + r$ to the type-0 child's investment and w_0 to the type-1 child's non-investment, so

$$\Delta V_L^0 = r. \tag{D.22}$$

The type-1 stake is nonnegative exactly when

$$q \geq q^{**}(P).$$

Thus the same condition that supports the low-education equilibrium also makes transmission of neighborhood orientation privately valuable for type-1 parents.

The low branch therefore changes transmission incentives. A type-1 child retains the full local-network payoff and conforms to the low local norm; a type-0 child invests and gives up part of that payoff. Higher r works in the opposite direction: it lowers ΔV_L^1 and raises ΔV_L^0.

In the minimal model,

$$\Phi_H(q) = 0. \tag{D.23}$$

This is not a claim that neighborhood orientation stops being transmitted in educated neighborhoods. It follows because education is the only payoff-relevant behavior through which type matters in (D.2). A richer specification with direct relational behavior could instead yield

$$\Phi_L(q) > \Phi_H(q) \geq 0.$$

The mechanism used below only requires transmission incentives to be stronger on the low branch.

Equation (D.23) also refers only to endogenous transmission drift. With residential turnover, q can continue to move on the high branch.

D.7 Dynamics of neighborhood orientation

Substituting (D.21)–(D.22) into (D.18), the low-branch flow is

$$\Phi_L(q) = \frac{q(1-q)}{\psi} g(q), \tag{D.24}$$

where

$$\begin{aligned} g(q) &= (1-q)(s_\Gamma q - \kappa - r) - qr \\ &= -s_\Gamma q^2 + (s_\Gamma + \kappa)q - (\kappa + r). \end{aligned} \tag{D.25}$$

Define

$$D \equiv (s_\Gamma + \kappa)^2 - 4s_\Gamma(\kappa + r). \tag{D.26}$$

Suppose

$$D > 0 \qquad \text{and} \qquad s_\Gamma > \kappa + 2r. \tag{D.27}$$

Then g has two interior roots,

$$q_c = \frac{s_\Gamma + \kappa - \sqrt{D}}{2s_\Gamma}, \qquad q^\dagger = \frac{s_\Gamma + \kappa + \sqrt{D}}{2s_\Gamma}, \tag{D.28}$$

with

$$q^{**}(P) < q_c < q^\dagger < 1. \tag{D.29}$$

Because g is a downward-opening quadratic,

$$\Phi_L(q)\begin{cases} <0, & q \in (q^{**}, q_c), \\ >0, & q \in (q_c, q^\dagger), \\ <0, & q \in (q^\dagger, 1). \end{cases} \tag{D.30}$$

The lower root q_c is a tipping point. The low-education equilibrium can exist for $q \geq q^{**}$ while neighborhood orientation still declines. Accumulation requires the stronger condition

$$q > q_c.$$

At the upper root,

$$g'(q^\dagger) < 0.$$

For the discrete-time map

$$q_{t+1} = q_t + \Phi_L(q_t),$$

local stability requires

$$-2 < \Phi_L'(q^\dagger) < 0. \tag{D.31}$$

The condition at the point is not enough for convergence from every initial state in $(q_c, 1)$; the discrete map could overshoot elsewhere. A global step-size bound closes this gap. Since $\Phi_L' = [(1-2q)g + q(1-q)g']/\psi$, with $|g| \leq \kappa + r + (s_\Gamma + \kappa)^2/4s_\Gamma$, $|g'| \leq s_\Gamma + \kappa$, and $q(1-q) \leq \frac{1}{4}$ on $[0,1]$,

$$\psi \;\geq\; \bar{\psi} \equiv \kappa + r + \frac{(s_\Gamma + \kappa)^2}{4s_\Gamma} + \frac{s_\Gamma + \kappa}{4} \tag{D.32}$$

implies $|\Phi_L'| \leq 1$ on all of $[0,1]$, so the map $q \mapsto q + \Phi_L(q)$ is increasing; the same bound keeps both efforts in (D.17) interior. An increasing map with fixed points $q_c < q^\dagger$ moves any initial condition in $(q_c, 1)$ that remains on the low branch monotonically toward $q^\dagger$. Under (D.32), which I maintain below, $q^\dagger$ is therefore stable with basin $(q_c, 1)$ on the low branch and q_c is unstable; $q = 1$ is a separate boundary fixed point because $q(1-q) = 0$.

The material payoff r shifts both roots. Since

$$\frac{\partial g}{\partial r} = -1,$$

the implicit-function theorem gives

$$\frac{dq^\dagger}{dr} = \frac{1}{g'(q^\dagger)} < 0, \qquad \frac{dq_c}{dr} = \frac{1}{g'(q_c)} > 0. \tag{D.33}$$

Using

$$\frac{dr}{dP} = -C'(P) > 0,$$

$$\frac{dq^\dagger}{dP} < 0, \qquad \frac{dq_c}{dP} > 0. \tag{D.34}$$

Poorer provision therefore raises the low-branch rest point and lowers the tipping point through the cost channel.

Starting from identical composition $\bar{q}$, divergence is conditional. If A is on the low branch and

$$q_c(r_A) < \bar{q} < q^{\dagger}(r_A), \tag{D.35}$$

then q_A rises toward $q^{\dagger}(r_A)$. If B remains on the high branch, $\Phi_H = 0$ and its composition stays at $\bar{q}$ in the absence of churn. If instead $\bar{q} \leq q_c(r_A)$, the education difference may remain while neighborhood orientation does not accumulate.

This is the feedback used in the model:

$$e \longrightarrow q$$

because the education regime changes transmission incentives, and

$$q \longrightarrow e$$

because a larger q raises the neighborhood pull. Neither direction is separately identified in the empirical analysis.

D.8 Convergence of provision

At date T, let the provision-related cost wedge disappear in the model:

$$C(P_A) = C(P_B) = C(\bar{P}), \qquad \bar{r} \equiv \rho - C(\bar{P}) > 0. \tag{D.36}$$

I assume

$$\bar{q} < q^{**}(\bar{P}), \tag{D.37}$$

so the city composition lies in the unique high-education region.

The low-education equilibrium remains available in A after convergence if

$$q_A(T) \geq q^{**}(\bar{P}), \tag{D.38}$$

and is the unique education equilibrium there if

$$q_A(T) > q^{*}(\bar{P}). \tag{D.39}$$

Persistence is therefore conditional on the inherited state. If the exclusion era is too short or the accumulated stock too small, provision convergence removes the low branch.

Convergence also changes the transmission dynamics, under two maintained assumptions. First, the modern standard improves on era provision in A: $C(\bar{P}) < C(P_A)$, so $\bar{r} > r_A$. Equalization across neighborhoods does not imply this by itself. Second, (D.27) continues to hold at $\bar{r}$, so the roots exist at the new payoff. Then

$$q^{\dagger}(\bar{r}) < q^{\dagger}(r_A), \qquad q_c(\bar{r}) > q_c(r_A). \tag{D.40}$$

The provision change therefore erodes part of the inherited difference even without residential mixing. If (D.27) instead fails at $\bar{r}$, the flow is negative on the whole low branch and convergence removes the interior rest point altogether; this strengthens erosion, and the persistence analysis below presupposes (D.27) at $\bar{r}$.

This distinction helps interpret the education outcomes in Section 3. Completed schooling is a stock formed across cohorts, including cohorts exposed to the historical provision asymmetry. Current attendance is a contemporary choice. The available contemporary provision measures show no discontinuity in school supply or quality at the boundary, so the model does not attribute the current attendance gap to the same measured provision wedge faced by older cohorts.

D.9 Residential churn

After T, let a fraction $\mu_\ell \in [0, 1)$ of residents each period be replaced by draws from the city distribution, whose type-1 share is $\bar{q}$. Then

$$q_{\ell,t+1} - q_{\ell,t} = (1 - \mu_\ell)\Phi_r(q_{\ell,t}) + \mu_\ell(\bar{q} - q_{\ell,t}). \tag{D.41}$$

On the high-education branch, $\Phi_H = 0$, so

$$q_{\ell,t+1} - q_{\ell,t} = \mu_\ell(\bar{q} - q_{\ell,t}). \tag{D.42}$$

There is no endogenous transmission drift, but residential turnover moves local composition toward the city mean.

On the low branch define

$$G(q, \mu) \equiv (1 - \mu)\Phi_L(q) + \mu(\bar{q} - q). \tag{D.43}$$

Suppose $q > \bar{q}$ is an interior stable fixed point:

$$G(q, \mu) = 0, \qquad G_q(q, \mu) < 0.$$

Stability refers to the discrete map $q_{t+1} = q_t + G(q_t, \mu)$, which requires $-2 < G_q < 0$. Under (D.32), $G_q = (1 - \mu)\Phi_L' - \mu \geq -1$, so the sign condition is also sufficient and the definition of $\bar{\mu}$ below is unaffected. At such a fixed point,

$$\begin{aligned} G_\mu &= -\Phi_L(q) + (\bar{q} - q) \\ &= -\frac{q - \bar{q}}{1 - \mu} < 0, \end{aligned} \tag{D.44}$$

where the second line uses $G(q, \mu) = 0$. Hence

$$\frac{dq}{d\mu} = -\frac{G_\mu}{G_q} < 0. \tag{D.45}$$

Higher churn therefore lowers any stable persistent fixed point while that fixed point exists.

For sufficiently large μ, the churn term dominates the positive part of Φ_L and no fixed point above $\bar{q}$ survives. Define

$$\bar{\mu} \equiv \sup \left\{ \mu \in [0,1) : \exists\, q > \bar{q} \text{ such that } G(q,\mu) = 0,\ G_q(q,\mu) < 0 \right\}. \tag{D.46}$$

For $\mu < \bar{\mu}$, a stable high-q configuration may persist. For $\mu > \bar{\mu}$, the local state moves downward until it leaves the low-education region; once

$$q < q^{**}(\bar{P}),$$

the low branch is unavailable. The subsequent high-education dynamics are given by (D.42), which sends q toward $\bar{q}$.

At a regular boundary value $\bar{\mu}$, disappearance of the stable and unstable interior fixed points occurs by tangency,

$$G(q,\bar{\mu}) = 0, \qquad G_q(q,\bar{\mu}) = 0, \tag{D.47}$$

the usual saddle-node case. I do not require uniqueness of this tangency for the persistence result.

The model therefore gives a simple race between local transmission and population turnover. Slow churn can preserve a locally different state; sufficiently rapid churn eliminates it.

Remark 1 (Residential rootedness and the 1992 episode). Type-1 agents on the low branch lose part of the local-network payoff when they leave the neighborhood. This gives them an additional reason to remain, a margin that could be modeled explicitly through residential choice as in Bezin and Moizeau (2017). It is consistent with the greater residential rootedness measured on the eastern side, but the model does not identify rootedness as the source of persistence.

A temporary increase in effective turnover pushes q_A toward $\bar{q}$. Whether the local state subsequently recovers depends on where the shock leaves it relative to the basin of the persistent fixed point. The 1992 explosions exercise in Section 5.5 is suggestive in this respect, but it bundles destruction, displacement, reconstruction, housing changes, and repopulation. It does not identify μ_A or the churn mechanism.

D.10 Extension: provision and the value of local networks

The baseline holds Γ fixed. Provision affects schooling through $C(P)$ only. As an extension, suppose

$$\Gamma = \Gamma(P), \qquad \Gamma'(P) < 0. \tag{D.48}$$

Formal provision may substitute for services supplied through local ties, so weaker provision raises the value of those ties. This channel is not identified in the empirical analysis.

The education condition becomes

$$\rho - C(P_\ell) \geq a\left[\delta\Gamma(P_\ell) q_\ell + \kappa(1 - 2\bar{e}_\ell)\right]. \tag{D.49}$$

The low-education threshold is

$$q^{**}(P) = \frac{\rho - C(P) + \kappa}{\delta\Gamma(P) + 2\kappa}, \tag{D.50}$$

with

$$\frac{dq^{**}}{dP} = \frac{-C'(P)}{\delta\Gamma(P) + 2\kappa} - \frac{[\rho - C(P) + \kappa]\delta\Gamma'(P)}{[\delta\Gamma(P) + 2\kappa]^2} > 0. \tag{D.51}$$

The first term is the baseline cost channel. The second is the network-substitution channel. Both make the low-education equilibrium harder to sustain as provision rises.

The two versions differ most clearly in their impact effect on the neighborhood pull. Holding $(q, \bar{e})$ fixed,

$$\left.\frac{\partial \Pi}{\partial P}\right|_{\text{baseline}} = 0, \tag{D.52}$$

whereas under (D.48),

$$\frac{\partial \Pi}{\partial P} = \delta q\, \Gamma'(P) < 0. \tag{D.53}$$

Thus lower provision raises the pull directly only under the extension. Provision already affects transmission in the baseline through $r = \rho - C(P)$; allowing Γ to depend on P adds a second channel by changing the value of local ties themselves.

The extension is useful for distinguishing mechanisms, but none of the persistence results requires it.

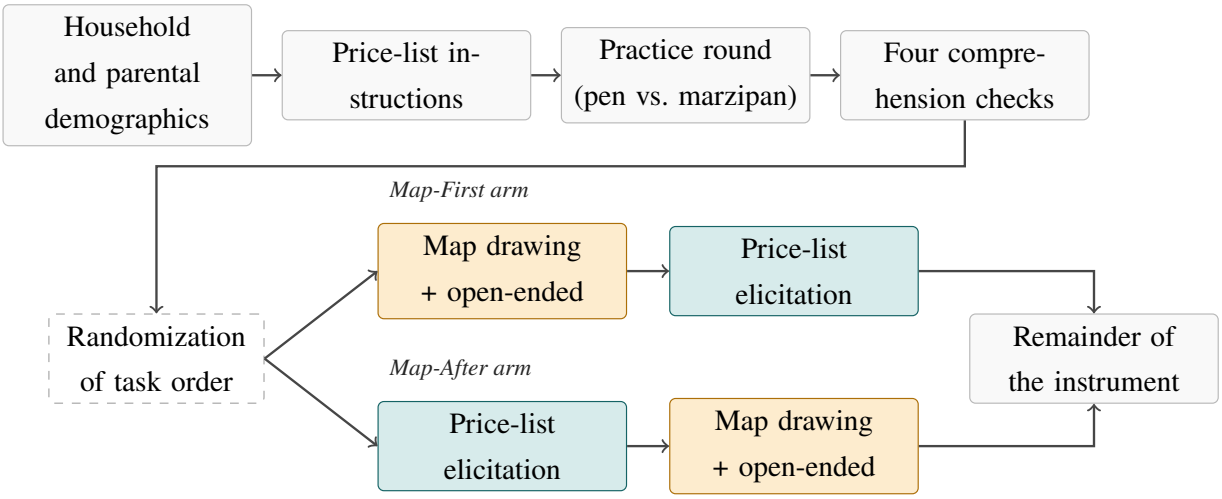
D.11 Empirical interpretation.

The model is not a decomposition of the reduced-form estimates. The spatial RDD identifies the long-run effect of the historical assignment under continuity, while the experiment identifies the effect of randomized neighborhood salience on educational demand. The historical and contemporary relational measures motivate the state q_ℓ, but they do not identify either direction of the feedback between q_ℓ and education. Likewise, neither the provision-to-orientation channel nor $\Gamma'(P) < 0$ is separately identified.

E The Incentivized Choice Experiment: Supporting Material

This appendix collects the supporting material for the incentivized choice experiment of Section 5, in the order the section draws on it. Figure E.1 sets out the order of tasks in the survey instrument. Appendix E.1 then reports the confirmatory table. Appendix E.2 characterizes what the priming activates, and Appendix E.3 reports balance, implementation diagnostics, and influential-observation checks. Appendix E.4 addresses experimenter-demand effects, Appendix E.5 the exercise that makes the historical divide salient, and Appendix E.6 details the survey instrument and protocols.

Figure E.1: Order of tasks in the incentivized choice experiment



Notes: Task order in the survey instrument. The price-list instructions, the practice round with a pen and marzipan candies, and the four comprehension checks precede the randomization and are therefore identical in both arms. The randomization determines only the order of the two remaining blocks: in the Map-First arm respondents draw their neighborhood and answer the open-ended question about the boundary they drew, and then complete the price list; in the Map-After arm they complete the price list first. Nothing unrelated to the neighborhood intervenes between the map-drawing module and the price list in the Map-First arm.

E.1 Confirmatory results

Table E.1 reports the full regression table behind the main-text figure: both outcomes, the pooled and interacted specifications, and the Analco subsample.

The confirmatory estimates of H1 (average effect) and H2 (heterogeneity by side) are reported in the main text, Table E.1.

Latent willingness to pay. The pre-registered willingness-to-pay measure (Table E.1) is the two-step latent estimator of Jack et al. (2022), registered as a primary outcome in AEARCTR-0018600. Table E.2 sets it against the area under each respondent's step demand curve, bounded at the $250 top of the price grid. The latent estimator works in two steps: I first recover each participant's *latent* willingness to pay from a scaled random-effects probit of their eleven binary choices on the card value — which accommodates never- and multiple-switching and, unlike the

Table E.1: Effect of neighborhood-salience priming on parents' educational investment

	Share of tutoring choices			Willingness to pay (MXN pesos)		
	OLS			Jack et al. (2022), two-step latent WTP		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{Map\text{-}First_i\}$	−0.077*	0.144	−0.117**	−44.7	97.6	−70.9**
Robust	(0.049)	(0.126)	(0.053)	(36.8)	(93.4)	(39.6)
Randomization inference	(0.046)	(0.117)	(0.051)	(30.7)	(80.8)	(33.4)
$\mathbb{1}\{Analco_i\}$		0.006			−18.4	
Robust		(0.219)			(172.7)	
Randomization inference		—			—	
$\mathbb{1}\{Map\text{-}First_i\} \times \mathbb{1}\{Analco_i\}$		−0.267**			−172.2*	
Robust		(0.137)			(102.9)	
Randomization inference		(0.129)			(88.4)	
Sample	Full	Full	East bank	Full	Full	East bank
Enumerator FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
LASSO controls	✓	✓	✓	✓	✓	✓
Distance to boundary	✓	✓	✓	✓	✓	✓
p-value: east-bank effect = west-bank effect (robust / RI)		0.052 / 0.037			0.094 / 0.051	
Dep. var. mean	0.548	0.548	0.538	164.0	164.0	155.6
Control mean (Map-After)	0.561	0.561	0.573	171.6	171.6	176.3
Control SD (Map-After)	0.384	0.384	0.372	255.9	255.9	249.5
Participants	239	239	193	239	239	193
Within R^2	0.235	0.251	0.216	0.222	0.238	0.208

Notes: Estimates from the pre-registered incentivized choice experiment near the historical boundary (AEA RCT Registry AEARCTR-0018600). *Map-First* equals one when the map-drawing module precedes the incentivized tutoring choice and zero when it follows; $\mathbb{1}\{Analco_i\}$ identifies households east of the historical boundary. Columns (1)–(3) report the share of eleven multiple-price-list rows in which the parent chose one hour of online math tutoring over an Amazon gift card. Columns (4)–(6) report latent willingness to pay (WTP) in MXN, estimated following the two-step procedure of Jack et al. (2022): a random-effects probit recovers individual posterior-mean WTP from the eleven choices, and the resulting WTP is regressed on the treatment indicators. All specifications include enumerator and day fixed effects, side-specific linear slopes in distance to the boundary, and controls selected by LASSO. Columns (1) and (4) report the pooled treatment effect; columns (2) and (5) allow it to differ by side, with the reported p-value testing equality of the two effects; columns (3) and (6) restrict the sample to Analco. The *Robust* rows report standard errors in parentheses, using HC1 with a finite-sample correction for the share outcome and a subject-level Bayesian bootstrap with 1,000 replications (Rubin, 1981) for WTP. The *Randomization inference* rows report the standard deviation of the estimate across 50,000 permutations of treatment within the original randomization strata, holding the number treated on each side fixed. The equality test imposes a constant common effect before permutation. No randomization-inference statistic is reported for $\mathbb{1}\{Analco_i\}$ because side is not randomized. Significance stars use randomization-inference p-values where available and robust p-values otherwise. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

bounded area, is not censored at $250 — and then regress that individual latent WTP on the treatment. The latent measure correlates 0.99 with the bounded area measure, so the two order participants almost identically; the latent effects are larger because the bounded scale is censored at the top of the grid. Both deliver the same conclusion: a negative priming effect on willingness to pay concentrated in Analco.

Table E.2: Willingness-to-pay treatment effects: bounded area vs. latent scaled-probit measure

	Pooled (H1)	Map-First × Analco (H2)	Effect in Analco
Latent scaled-probit WTP (two-step, main text)	−44.7 (0.146)	−172.2 (0.051)	−74.6 (0.025)
Area under demand ($0–$250, participant RE)	−18.4 (0.130)	−67.7 (0.047)	−30.2 (0.022)

Notes: The first row reproduces the two-step latent willingness-to-pay effects of Table E.1; the second row re-estimates them with the bounded area-under-demand measure. The latent estimator is a two-step procedure following Jack et al. (2022): step one recovers each participant's latent willingness to pay (in pesos) as the individual posterior mean from a scaled random-effects probit of the eleven binary choices on the card value (with row- and column-order controls) estimated *without* the treatment; step two regresses that individual WTP on the Map-First indicator and its Analco interaction. The area measure is computed as the area under the step demand curve over the ten positive-price rows, so it is bounded at 250 pesos and estimated with a participant-level random intercept and CR2 standard errors clustered on the participant. The latent measure correlates 0.991 with the bounded area measure — they order participants almost identically — but is un-censored, so its scale (and hence the effects) is wider. Two-sided p-values are in parentheses. They come from a subject-level Bayesian bootstrap with 1,000 replications that reweights subjects and repeats both steps, which corrects for the generated-regressand problem. The effect in Analco is here the sum of the Map-First coefficient and its Analco interaction, so it differs slightly from the east-bank subsample estimate reported in Table E.1 and Table E.7. Both measures deliver the same conclusion: a negative priming effect concentrated in Analco.

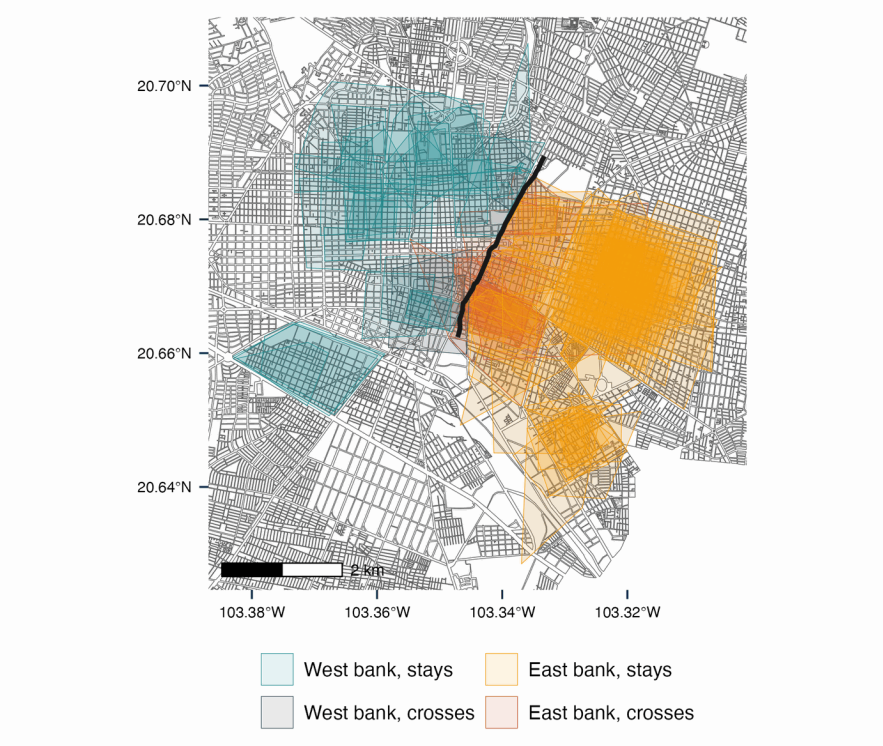
E.2 Opening the black box: what the neighborhood priming activates

A standard objection to priming manipulations is that they are "black boxes": they change behavior without revealing which mental representations they activate (Cohn and Maréchal, 2016). The map-drawing module is unusually transparent on this point, because the object the respondent produces — a hand-traced neighborhood — *is* the activated representation. I characterize it in two ways: the geometry of the drawn barrio and what respondents say they had in mind. Both are descriptive. Both arms draw and answer immediately afterward, so a contrast between them would reflect what preceded the drawing rather than the drawing itself.

The geometry of the drawn barrio. Table E.3 summarizes the polygons. Each traced barrio is represented by the convex hull of its vertices. The neighborhoods people draw are small (a median of just over half a square kilometer — a few blocks), and home-centered: for 88% of respondents the geolocated home falls *inside* their own drawing. Only 11% of the drawn barrios cross the historical boundary, the SJdD creek, today Calzada Independencia (Figure E.2). That

figure should not be read as avoidance of the line, because most respondents live too far from it for a drawing of this size to reach it. A polygon of the median area corresponds to a circle of about 430 metres in radius, while the median distance from home to the boundary is close to a kilometre, so a home-centred drawing of typical size falls several hundred metres short of the avenue before any question of identity arises. The exercise reports no opportunity-to-cross adjustment and no placebo line elsewhere in the city against which to compare the rate, and a drawing that stops at a major avenue may simply be stopping at a salient physical landmark. What the polygons establish is narrower and still useful: when asked for their neighborhood, people draw a compact area around their own home rather than a district or a side of the city. The table reports the geometry overall and by side of the boundary.

Figure E.2: Neighborhoods traced by respondents, relative to the historical boundary



Notes: Each polygon represents the neighborhood delineated by one respondent in the interactive mapping exercise, drawn as the convex hull of the traced vertices. The dark line denotes the historical San Juan de Dios creek boundary, now Calzada Independencia. Polygons are colored by the respondent's own side of the historical boundary and by whether the traced neighborhood crosses it: teal for west-bank respondents whose drawing stays on their side and navy for those whose drawing crosses, amber for east-bank respondents whose drawing stays and brick for those whose drawing crosses. Polygons with obvious drawing errors, such as a misplaced distant vertex, are omitted.

What respondents said they thought about. The open-ended companion question describes the same object. I classify the free-text answers with `claude-opus-4-8` against a fixed codebook (Section E.6.1), and a second, independent classification of the same answers, formed without access to the first, agrees closely: mean Cohen's $\kappa = 0.93$ across the six themes, and 0.98 for social ties. Table E.4 reports the shares by side of the boundary.

Table E.3: Geometry of the neighborhood respondents drew

	All	Analco	West bank
Barrio area (km^2), median	0.58	0.56	0.71
Barrio area (km^2), mean	0.79	0.77	0.91
Home inside the drawn barrio	0.876	0.862	0.933
Home off-centering (dist./$\sqrt{\text{area}}$)	0.40	0.39	0.43
Crosses the historical boundary	0.112	0.122	0.067
Vertices drawn	4.3	4.3	4.4
Participants	233	188	45

Notes: Geometry of the neighborhood each respondent traced on the interactive map, for the 233 of 239 participants who drew a valid polygon of at least three vertices (3 far-vertex drawing errors, with a vertex over 5 km from the home, are dropped). Each traced neighborhood is represented by the convex hull of its vertices, so that self-intersecting traces still define a well-behaved convex region. The barrio area is that hull's area, reported as both the median and the mean because the distribution is right-skewed; the remaining rows are means. *Home inside* is the share whose geolocated home falls within their own drawing; *home off-centering* is the distance from the home to the polygon centroid divided by the square root of the area, so lower values are more centered; *crosses the historical boundary* is the share whose drawing intersects the SJdD creek, today Calzada Independencia. The drawings are small, home-centered, and rarely cross the historical boundary. The table is descriptive. Both experimental arms draw their neighborhood and answer immediately afterward, so treatment-arm comparisons are not reported.

The neighborhood is conceived as a concrete place that respondents walk and use. Physical extent is the most common theme, at 62%: streets, avenues, landmarks, and how far the area reaches. Everyday use follows at 49%, pooling lived familiarity and daily routine ("where I move," "what I know," "as far as I can walk") with errands, shops, work and school. People and belonging appear in 13% of answers and safety in 6%, with social ties named by one respondent in ten. The mapping task therefore brings a place to mind more than it brings a community to mind, which is what the polygons also show. That is the object urban sociology describes, a neighborhood constituted by routine interaction grounded in shared residence and the common use of local facilities rather than by dense, intimate bonds (Smith, 2010; Sampson et al., 2002), as Section 5.2 sets out. What the exercise does not support is a claim about local identity.

The table reports no test. The experimental arm is absent by design: both arms answer this question immediately after drawing, so a contrast between them would identify the effect of the preceding multiple-price list on the description rather than what the mapping task brings to mind. Differences by side are small, and none of the six themes is distinguishable from zero once the multiplicity of the comparisons is taken into account.

Table E.4: What respondents had in mind when drawing their neighborhood

	All	Analco	West bank
Physical extent	0.62	0.63	0.61
Everyday use	0.49	0.50	0.48
Daily routine or familiarity	0.37	0.37	0.39
Errands, shops, work, or school	0.19	0.20	0.13
People or belonging	0.13	0.15	0.07
Social ties	0.10	0.11	0.04
Long residence or family roots	0.04	0.04	0.02
Safety or calm	0.06	0.05	0.11
Participants	239	193	46

Notes: Share of respondents whose open-ended answer to "what did you have in mind when drawing the limits of your neighborhood?" mentions each theme. Themes are coded by `claude-opus-4-8` using a fixed codebook and are not mutually exclusive. *Physical extent* includes streets, avenues, landmarks, and the reach of the area; *Everyday use* includes daily routines and places respondents regularly use; *People or belonging* includes social ties, long residence, and family roots. An independent classification yields mean Cohen's $\kappa = 0.93$ across themes. The table is descriptive. Both experimental arms answer this question immediately after drawing their neighborhood, so treatment-arm comparisons are not reported.

E.3 Balance, diagnostics, and influential observations

Covariate balance. Table E.5 reports balance across the Map-First (treatment) and Map-After (control) arms, in two panels. Panel A covers the twenty-six variables recorded before the randomized block — side of the boundary and orthogonal distance to it, parent and grandparent education, household composition, income bracket, and price-list comprehension — and none differs significantly across arms, individually or after a sharpened false-discovery-rate correction (Anderson, 2008). Side of the boundary is the randomization stratum: it is identical under every admissible permutation and therefore carries no test. Panel B reports the focal child's characteristics and the stated reason for choosing that child. The household roster is recorded before treatment, but which child becomes the focal one is decided afterwards, so these are not balance variables and the treatment could in principle have shifted the choice. None of the six differs across arms either, so parents in the two arms are making the tutoring decision about comparable children.

Implementation diagnostics. Table E.6 reports response-quality diagnostics by arm. Comprehension is high — 87% of respondents answer all four price-list comprehension questions correctly — and none of the quality metrics differs across arms, so the priming manipulation did not degrade response quality; treatment assignment is also orthogonal to enumerator identity ($F = 1.45$, $p = 0.21$). As is common in the multiple-price-list literature (Jack et al., 2022), a non-trivial share of respondents exhibit *never-switching* behaviour (43% choose the same option in all eleven MPL rows) or *multiple-switching* behaviour (22% switch more than once); Figure E.3 plots the distribution of switches in the tutoring MPL separately for the treatment and control arms, which look alike — the manipulation did not affect elicitation quality. This is why non-monotonic switching is treated as a comprehension flag rather than a discarded observation: the pre-registered latent estimator of Jack et al. (2022) accommodates never- and multiple-switchers, as does the bounded area under each respondent's demand curve reported alongside it below.

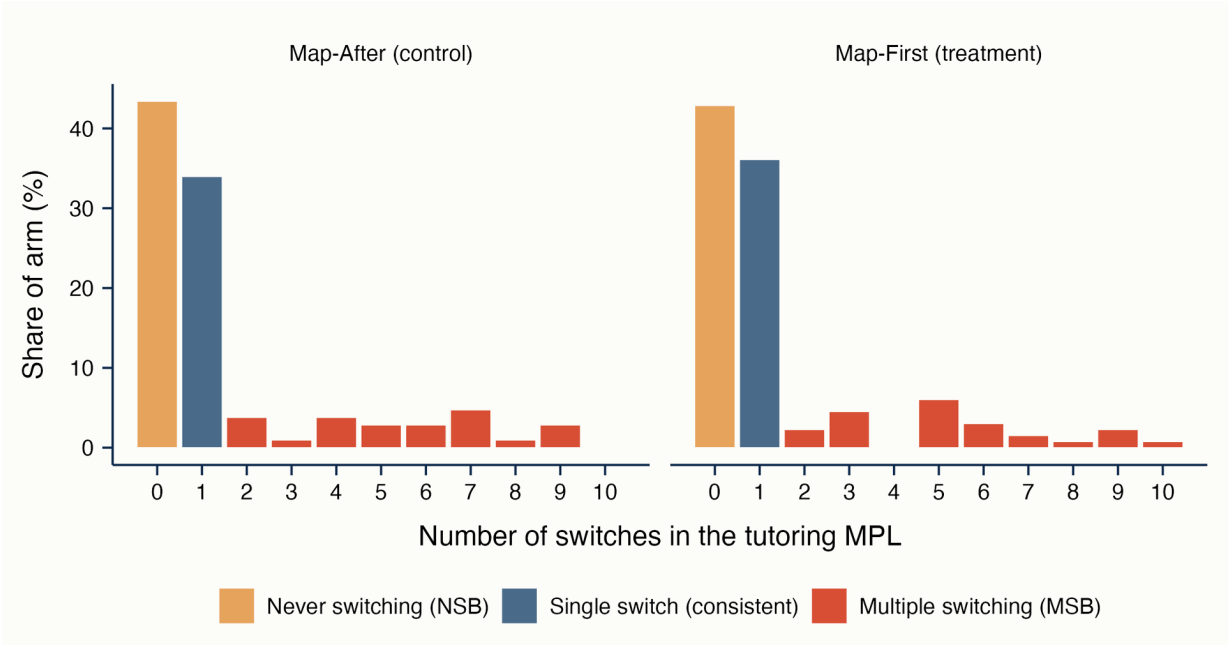
Randomization inference and influential observations. Tables E.1 and E.7 probe the two pre-registered results against the small sample and the possibility that a few households drive them. Randomization inference is reported in Table E.1, which permutes treatment 50,000 times within the original randomization strata, holding the number treated on each side fixed; the significance stars in that table are computed from those permutations. The Map-First × Analco interaction (H2) is significant on both outcomes, with randomization-inference p-values of 0.037 for the tutoring share and 0.051 for willingness to pay, and the implied priming effect in Analco is significant at the 5 percent level on both. The pooled effect (H1) is significant at the 10 percent level for the tutoring share and is not significant for willingness to pay. Table E.7 re-estimates each effect dropping one participant at a time. The Analco priming effect stays within a narrow band that never approaches zero, $[-0.130, -0.103]$ for the tutoring share and $[-78.0, -62.3]$

Table E.5: Covariate balance across the Map-First treatment

	Map-After (control)	Map-First (treatment)	Difference (T − C)	p-value	q-value	N
Panel A. Measured before the randomized block						
$\mathbb{1}\{Analco_i\}$	0.840	0.782	-0.058	—	—	239
Distance to boundary (m)	1012	959	-53	0.902	1.000	239
Parent age	44.93	43.53	-1.40	0.327	1.000	239
Female respondent	0.764	0.699	-0.065	0.302	1.000	239
Parent completed high school	0.594	0.586	-0.008	1.000	1.000	239
Father completed high school	0.344	0.319	-0.026	0.766	1.000	203
Mother completed high school	0.265	0.283	0.019	0.767	1.000	229
Has partner	0.660	0.684	0.024	0.783	1.000	239
Adults in household	2.87	2.82	-0.05	0.800	1.000	239
Minors in household	1.78	1.80	0.01	0.957	1.000	239
Rooms in dwelling	2.73	2.72	-0.00	1.000	1.000	239
Crowding (persons/room)	1.79	1.89	0.10	0.311	1.000	239
Children aged 6–20	1.66	1.65	-0.01	0.937	1.000	239
Respondent works	0.651	0.692	0.041	0.574	1.000	239
Income: under \$6,000	0.149	0.117	-0.032	0.541	1.000	214
Income: \$6,001–9,000	0.149	0.125	-0.024	0.689	1.000	214
Income: \$9,001–12,000	0.160	0.192	0.032	0.597	1.000	214
Income: \$12,001–15,000	0.128	0.167	0.039	0.446	1.000	214
Income: \$15,001–18,000	0.085	0.083	-0.002	1.000	1.000	214
Income: \$18,001–22,000	0.074	0.158	0.084	0.088	1.000	214
Income: \$22,001–27,500	0.106	0.050	-0.056	0.185	1.000	214
Income: \$27,501–35,000	0.085	0.075	-0.010	0.798	1.000	214
Income: \$35,001–55,000	0.053	0.017	-0.037	0.253	1.000	214
Income: over \$55,000	0.011	0.017	0.006	1.000	1.000	214
Main income is a formal salary	0.755	0.729	-0.025	0.771	1.000	239
Price-list comprehension (0–4)	3.81	3.86	0.05	0.423	1.000	239
Panel B. Measured after the randomized block						
Focal child female	0.425	0.414	-0.011	0.895	1.000	239
Focal child age	12.80	12.40	-0.40	0.447	1.000	239
Focal child attends school	0.896	0.925	0.029	0.502	1.000	239
Chose child: needs academic help	0.094	0.090	-0.004	1.000	1.000	239
Chose child: ability/aspiration	0.094	0.053	-0.042	0.304	1.000	239
Chose child: age/only eligible	0.321	0.429	0.108	0.105	1.000	239

Notes: Covariate balance across the Map-First (treatment) and Map-After (control) arms of the pre-registered incentivized choice experiment. Panel A reports variables recorded before the randomized block of the instrument, items (i) to (iv) of Section E.6.1, including the price-list comprehension checks, which are administered before assignment; these are the balance tests proper. Panel B reports the characteristics of the focal child and the stated reason for choosing that child, recorded at item (vi), after the randomized block. The household roster from which the focal child is drawn is collected before treatment, but the choice of which child becomes focal is not, so Panel B does not test randomization and instead asks

Figure E.3: Distribution of tutoring-MPL switches by treatment arm: never-switching and multiple-switching behaviour



Notes: Number of switches between the two options across the eleven rows of the incentivized tutoring multiple price list (MPL), following the diagnostics of Jack et al. (2022), shown as the share of each arm. *Never-switching* respondents (0 switches, amber) choose the same option throughout; a single switch (teal) is the consistent pattern; *multiple-switching* respondents (> 1 switch, red) make inconsistent choices. The left panel is the Map-After (control) arm and the right panel the Map-First (treatment) arm; the two distributions are similar, indicating the priming manipulation did not change elicitation quality.

Table E.6: Incentivized choice experiment: implementation diagnostics

	Map-After (control)	Map-First (treatment)	All	Difference p-value
Comprehension score (0–4)	3.81	3.86	3.84	0.423
Passed all comprehension checks	0.840	0.902	0.874	0.184
Consistent on the practice MPL	0.547	0.624	0.590	0.250
Never-switching (NSB)	0.434	0.429	0.431	1.000
Multiple-switching (MSB)	0.226	0.211	0.218	0.874
Takes the \$0 card at a zero price	0.179	0.203	0.192	0.752
Enumerators / field days	6 enumerators, 15 days; treatment $\perp$ enumerator: $F = 1.45$, $p = 0.215$			

Notes: Response-quality diagnostics for the incentivized choice experiment, by treatment arm. Rows report the mean of each indicator among control (Map-After) and treatment (Map-First) respondents and overall, with the randomization-inference p-value (Map-First permuted within side of the historical boundary, the randomization stratum). The comprehension score counts correct answers to the four price-list comprehension questions; "passed" requires all four. "Consistent on the practice MPL" flags respondents who switch at most once (and not into a dominated choice) on the pen-versus-candy practice list. Never-switching (NSB) respondents choose the same option in all eleven MPL rows; multiple-switching (MSB) respondents switch more than once; the last row reports the share who take the \$0 gift card when tutoring is offered at a price of zero. That choice is not treated as an error: reading it as dominated would require assuming that tutoring carries no non-monetary cost and weakly positive value for every household, and online tutoring involves time, scheduling and supervision, so a zero or negative valuation is coherent. The row is reported as a rate and no respondent is excluded on its basis. No metric differs across arms, and a joint test finds treatment assignment orthogonal to enumerator identity, indicating the priming manipulation did not degrade response quality.

pesos for willingness to pay, so no single household drives the result.

E.4 Experimenter-demand effects

A priming experiment invites the concern that respondents guess the hypothesis and answer accordingly. After the elicitation, an open-ended question (Q88) asks each respondent what they think the study is about; I classify these guesses with `claude-opus-4-8` against a fixed codebook (Section E.6.1) — the same procedure used for the map-drawing responses — coding whether each guess refers to the neighborhood, education, economic conditions, or security, and whether it connects the neighborhood to education ("guessed the mechanism"); the coding agrees closely with an independent keyword dictionary on the flags central to the demand check (Cohen's κ between 0.80 and 0.97 for the neighborhood, education, security, and guessed-mechanism indicators). Only a small minority — 30 of 239 (13%) — guess the mechanism. Panel B of Table E.8 regresses these indicators on the treatment. The Map-First coefficient on guessing the mechanism is +0.22 with a randomization-inference p-value of 0.12, so the manipulation does not significantly raise the inference, and the point estimate sits on the west bank rather than in Analco. The Map-First×Analco interaction is −0.22, which leaves the probability of guessing essentially unchanged on the excluded side, and the effect on merely mentioning the neighborhood is smaller and further from significance. Whatever rise there is does not track the

Table E.7: Robustness of the incentivized-choice-experiment estimates: leave-one-out

	Estimate	Leave-one-out range
Panel A. Share of tutoring choices		
Pooled effect (H1)	-0.077	[−0.090, −0.067]
Map-First × Analco (H2)	-0.267	[−0.297, −0.195]
Priming effect in Analco	-0.117	[−0.130, −0.103]
Panel B. Willingness to pay (MXN pesos)		
Pooled effect (H1)	-44.7	[−53.1, −38.5]
Map-First × Analco (H2)	-172.2	[−195.3, −127.8]
Priming effect in Analco	-70.9	[−78.0, −62.3]

Notes: Robustness of the two pre-registered incentivized-choice-experiment estimates. Each row is an estimand from the specifications of Table E.1: the pooled Map-First effect (H1), the Map-First × Analco interaction (H2), and the priming effect in Analco, estimated on the east-bank subsample. "Estimate" reproduces the point estimate; its inference, conventional and by randomization, is reported in Table E.1. The leave-one-out range is the minimum and maximum of the estimate across the 239 re-estimations that each drop one participant; each re-estimation repeats both steps of the latent estimator. Willingness to pay is the pre-registered two-step latent measure of Jack et al. (2022) reported in Table E.1. The three rows replicate columns (1)–(3) of that table: the pooled effect, the Map-First × Analco interaction, and the east-bank effect estimated on the Analco subsample. All specifications carry the enumerator and day fixed effects, the LASSO-selected controls, and the Analco-by-distance terms of the main table.

side on which the treatment moves behavior. Panel A re-estimates the Analco priming effect on both outcomes and on both willingness-to-pay measures, the pre-registered latent one and the bounded area. Controlling for the inferred themes gives −0.124 for the tutoring share, −75.3 pesos for latent willingness to pay and −30.6 for the bounded area; dropping the 30 guessers gives −0.124, −78.8 and −30.6. Against baselines of −0.117, −71.1 and −28.6, every estimate is unchanged or slightly stronger. These checks are reassuring rather than decisive. Question Q88 is asked after both arms have completed the tasks, so it records what respondents infer at the end of the interview and not what they inferred at the moment of choice. The inferred themes and guesser status are themselves measured after treatment, so controlling for them or selecting on them shows that the estimate is stable rather than establishing that the intention-to-treat effect is free of demand. What the exercise does support is that a simple experimenter-demand account does not fit the pattern across banks, since the guessing moves on the side where behavior does not.

E.5 The boundary identity and how schooling is explained

A second, exploratory experiment embedded in the same survey makes the historical boundary salient through the expression associated with it rather than through the neighborhood as a place, and asks what changes in the explanations respondents give. They explain, in their own words, why a young person from their colonia would not continue studying after high school (question Q91), and whether Q90 — how often they hear or use *de la Calzada para allá* — is asked before or after that question is randomized, without stratification. The randomization identifies the effect of making that expression salient. It does not identify what the expression carries, which may be identity, stigma, geography, or perceived inequality, and the exercise does not separate them or claim to have activated a neighborhood identity.

Each answer is classified against a fixed codebook into the reasons it gives: the local place, the surrounding social environment, the family, material constraints, and the young person's own choices. Answers typically give more than one reason, so the categories are not exclusive. A second, independent classification of the same answers, blind to the arm and formed without access to the first, agrees closely, and reproduces the estimate below.

Table E.9 reports the results. On the excluded side the prime displaces the material explanation: the share of Analco respondents attributing non-continuation to money or the need to work falls by 15 percentage points, which survives a false-discovery-rate correction over the five outcomes of the panel, the four categories and the derived neighborhood-minus-economic contrast. On the west bank the same displacement is 2 percentage points and indistinguishable from zero. Respondents give the same number of reasons in both arms, so the shift is not an artifact of shorter answers. The table reports the two banks separately and no cross-bank test. A displacement in Analco alongside a null on the west bank does not establish that the two sides differ, and with 46 west-bank respondents the design has little power to tell them apart.

Table E.8: Robustness to experimenter-demand effects

Panel A. Robustness of the priming effect in Analco

	(1)	(2)	(3)
Share of tutoring choices	−0.117**	−0.124**	−0.124**
[Map-First + Map-First×Analco]	(0.052)	(0.051)	(0.056)
Willingness to pay, latent (pesos)	−71.1**	−75.3**	−78.8**
[Map-First + Map-First×Analco]	(33.9)	(33.3)	(36.6)
Willingness to pay, bounded area (pesos)	−28.6**	−30.6**	−30.6**
[Map-First + Map-First×Analco]	(13.6)	(13.3)	(14.7)
Guess-theme controls		✓	
Sample	All	All	Non-guessers
Participants	239	239	209

Panel B. Effect of the treatment on the inferred topic (linear probability models)

	Map-First	Map-First × Analco	Dep. var. mean
Guessed mechanism	0.217	−0.224*	0.126
	(0.102)	(0.111)	
Mentioned neighborhood	0.178	−0.131	0.360
	(0.139)	(0.155)	
Participants		239	

Notes: The open-ended guesses (question Q88, "what do you think the study is about?") are classified with `claude-opus-4-8` at temperature 0, one independent call per response, against a fixed codebook; the structured outputs are archived with the replication package (`llm-coding/q88_opus_coded.csv`). *Panel A* reports the implied priming effect in Analco (the linear combination Map-First + Map-First × Analco) from the interaction specification of Table E.1, for the share of tutoring choices and for willingness to pay. Both willingness-to-pay measures are reported, the pre-registered latent two-step measure of Table E.1 and the bounded area under the respondent's demand curve, which is censored at the top of the price grid and therefore smaller in magnitude. Column (1) is the baseline; column (2) adds controls for the themes the guess mentions (neighborhood, education, economic conditions, security); column (3) drops respondents who guessed the mechanism. *Panel B* regresses, by linear probability model, an indicator for whether the guess named the neighborhood–education mechanism ("Guessed mechanism") or mentioned the neighborhood ("Mentioned neighborhood") on the Map-First treatment, the Analco indicator, and their interaction. Overall, 30 of 239 (13%) guess the mechanism. The treatment does not significantly raise the probability of guessing it, and the point estimate sits on the west bank rather than in Analco, while the Analco priming effect is unchanged when the inferred themes are controlled for or the guessers dropped. Q88 is asked after both arms have completed the tasks, and the inferred themes and guesser status are measured after treatment, so these columns show that the estimate is stable rather than establishing that it is free of experimenter demand. HC1 heteroskedasticity-robust standard errors in parentheses; significance is assessed by randomization inference, permuting the Map-First label within side of the historical boundary, the randomization stratum. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table E.9: Boundary salience and explanations of educational non-continuation

	Share mentioning		Prime effect		Adjusted	
	Not primed	Boundary primed	Effect	q	Effect	q
Panel A. Analco, historically excluded bank ($N = 193$)						
Economic	0.905	0.755	-0.150^{***}	0.020	-0.170^{***}	0.005
Neighborhood	0.084	0.143	0.059	0.350	0.067	0.170
Family	0.095	0.143	0.048	0.395	0.039	0.320
Individual	0.316	0.357	0.041	0.637	0.044	0.320
Neighborhood − economic			0.209^{***}	0.020	0.237^{***}	0.005
Panel B. West bank ($N = 46$)						
Economic	0.880	0.857	−0.023	1.000	−0.107	0.791
Neighborhood	0.200	0.143	−0.057	1.000	−0.129	0.791
Family	0.160	0.238	0.078	1.000	0.126	0.791
Individual	0.480	0.286	−0.194	1.000	−0.149	0.791
Neighborhood − economic			−0.034	1.000	−0.022	0.791

Notes: At the end of the survey, respondents explain in their own words why a young person from their *colonia* might not continue studying after high school. I randomize whether respondents are first asked how often they hear or use the expression "on the other side of the Calzada" (*de la Calzada para allá*), associated with the historical divide. *Boundary primed* denotes respondents who receive this item before the open-ended education question; *Not primed* denotes those who receive it afterward. The order is randomized without stratification.

Each answer is classified by `claude-opus-4-8` using a fixed codebook. Categories are not mutually exclusive because respondents can give multiple reasons. *Economic* includes money and the need to work. *Neighborhood* includes explanations referring to the local place or surrounding social environment, including friends and what other young people nearby do. *Family* and *Individual* capture explanations referring to the family and the young person's own choices, respectively.

The first two columns report treatment-arm means. The *Prime effect* columns report the difference between the Boundary-primed and Not-primed groups. The *Adjusted* columns report the same difference after partialling out the randomized Map-First assignment from the earlier module, the duration of the interview in minutes, and the respondent's sex, age and age squared. Q91 is asked at the end of the survey, so the adjustment absorbs both the earlier experimental arm and respondent fatigue; duration is winsorized at its 95th percentile because a few sessions were left open. Because the reference distribution is generated under the sharp null, the adjusted p-values are exact whether or not the adjustment model is correct: the controls move precision, not validity. Inference is by randomization inference using 50,000 permutations of question order, conducted within bank. The q columns report sharpened false-discovery-rate values (Anderson, 2008) within each panel, computed separately for the unadjusted and adjusted estimates. The family is the five rows of the panel, the four categories together with the derived neighborhood-minus-economic contrast. The last row is the within-respondent difference between the neighborhood and economic indicators, which takes the values −1, 0 and 1. It answers whether the prime shifts the relative weight of the two explanations rather than the level of either, so it is treated as an outcome in its own right and enters the false-discovery-rate correction; it is not a share, so its arm means are omitted.

Respondents give a similar number of reasons in the two treatment arms (0.68 versus 0.79). An independent classification of the same answers, conducted without access to the first, yields mean Cohen's $\kappa = 0.83$ across categories and $\kappa = 0.94$ for the economic category. The analysis is exploratory. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

E.6 Incentivized choice experiment: additional design detail

This subsection collects further detail on the incentivized choice experiment beyond Section 5.2.

E.6.1 Survey instrument

The full instrument runs approximately 35–42 minutes and includes, in order: (i) household and parental demographics; (ii) a child-by-child loop for each child aged 6–20; (iii) parental work and time use; (iv) the price-list instructions, the practice round, and the four comprehension checks, all administered before the randomization and therefore identical across arms; (v) the randomized block, in which the price-list elicitation and the interactive map-drawing module with its open-ended follow-up are administered in the order given by treatment assignment, with the 1-in-10 ex-post lottery resolved after the price list (Figure E.1); (vi) focal-child selection; (vii) subjective lifetime returns to schooling; (viii) aspirations and expectations; (ix) the misperceptions module (peer beliefs at the $100 and $200 reference prices); (x) personal and peer norms about work-versus-study; (xi) trust and participation items in JCV format; (xii) the Martínez–Ramírez vignettes; (xiii) network benefits (emergencies, loans, job-finding); (xiv) hours-with-neighbors; (xv) belonging and identity; (xvi) residential stability and roots, including 1992-explosion exposure; (xvii) preferences (qualitative GPS items); (xviii) disinterested dictator games and an altruism item; (xix) the demand-effects open-ended question; and (xx) the secondary priming experiment on *"de la Calzada para allá."* The complete instrument is reproduced in the supplementary materials.

E.6.2 Randomization protocol

All randomizations are implemented within SurveyToGo at survey start. Three independent randomizations are conducted: (1) the salience treatment (Map-First vs Map-After), a fair binary draw stratified through the sampling quotas; (2) the local-expression priming (secondary), an independent fair binary draw; and (3) survey length (core 35-minute vs extended 42-minute versions), with all primary outcomes present in both. Treatment assignment is logged at survey start and exported with the response data; enumerators learn the assignment only when the software prompts the corresponding module.

E.6.3 Price-list comprehension protocol

The protocol consists of (a) a six-row practice round (BIC pen vs. marzipan candies); (b) four comprehension checks covering the random-row mechanism, the binding choice, the inability to revise ex post, and the 1-in-10 disbursement; and (c) a re-explanation step if any check fails. Pilot results show 92% comprehension on the first pass and 100% after one round of clarification.

F External Validity: Supporting Material

This appendix collects the supporting material for Section 6, in the order the section draws on it. The localization of social ties is measured from the Facebook Social Connectedness Index (Bailey et al., 2018) as the share of a place's friendship links that stay within a short radius; the complementary distant-ties share is the fraction reaching far. Neither is a direct measure of neighborhood orientation, and both are one observable dimension of it. Table F.1 covers the world's subnational regions (schooling from the Global Data Lab), Table F.2 the NUTS2 regions of Europe (tertiary attainment from Eurostat, the developed-economy counterpart absent from the Global Data Lab sample), and Table F.3 the municipalities of Mexico (2020 Census), with the scope condition estimated against, respectively, ethnic segregation and block-level SES segregation. Table F.4 reports U.S. ZIP-code areas (ACS 5-year), where county residential segregation moderates the gradient. Appendix F.1 details the World Values Survey sample behind the individual-level evidence.

Table F.1: Localization of social ties and schooling across subnational regions

	Mean yrs. 20+				Mean yrs. 20–39				Expected yrs.			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Localization of social ties (share ≤50 km)	-2.650***	-2.806***	-1.706***	-3.130***	-2.442***	-2.583***	-1.473***	-2.752***	-1.776***	-1.880***	-1.173***	-2.292***
	(0.2357)	(0.3497)	(0.3540)	(0.4873)	(0.2348)	(0.3481)	(0.3686)	(0.5115)	(0.2577)	(0.4049)	(0.4396)	(0.5432)
	[0.3394]	[0.4515]	[0.4381]	[0.5399]	[0.3471]	[0.4369]	[0.4292]	[0.5657]	[0.3685]	[0.4597]	[0.4641]	[0.5600]
Distant social ties (share ≥500 km)		-0.3266	0.7916	-0.2984		-0.2947	0.7895	-0.1324		-0.2171	0.3383	-0.4706
		(0.5701)	(0.5105)	(0.6426)		(0.5616)	(0.5320)	(0.6867)		(0.5920)	(0.5764)	(0.6643)
		[0.9035]	[0.7712]	[0.8044]		[0.7963]	[0.7166]	[0.7237]		[0.8018]	[0.7645]	[0.7150]
Localization of social ties × Ethnic segregation				-0.4000*				-0.3980				0.0820
				(0.2398)				(0.2757)				(0.2842)
				[0.2639]				[0.3104]				[0.3222]
Population			✓	✓			✓	✓			✓	✓
Area			✓	✓			✓	✓			✓	✓
Observations	1,145	1,145	901	474	1,145	1,145	901	474	1,080	1,080	836	441
R^2	0.840	0.840	0.895	0.891	0.843	0.843	0.897	0.893	0.734	0.734	0.808	0.830
Within R^2	0.134	0.135	0.311	0.386	0.111	0.111	0.283	0.351	0.055	0.055	0.214	0.282
Country fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Region-level (GDL) regressions with country fixed effects. HC1 robust standard errors appear in parentheses and, below them in brackets, standard errors clustered on country. The segregation moderator varies only across countries, so the clustered errors are the relevant ones for column (iv); under them the interaction is not distinguishable from zero for any of the three outcomes, and the moderation result of Section 6 rests on Table F.4 rather than on this table. Localization of social ties = share of a region's Facebook friendship links to places within 50 km; distant ties = share to places 500 km or farther (haversine between GADM1 centroids). Cols. (i)-(iv) per outcome: (i) localization only; (ii) + distant ties; (iii) + controls (log population, log area); (iv) + interaction with ethnic segregation (Alesina-Zhuravskaya). The segregation index is unavailable for many countries, so column (iv) is estimated on 474 regions against 901 in column (iii), and the two columns are not comparable. Re-estimating column (iii) on the 474-region sample gives −3.230 (0.488) for mean years at 20+, against −1.706 (0.354) on the full sample and −3.130 (0.487) in column (iv), so the change between columns (iii) and (iv) is a change of sample rather than of specification. The localization coefficient in column (iv) is the slope at zero standardized segregation, not the average slope. Standard errors in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

F.1 The World Values Survey sample

Table F.5 reports the countries available for the individual-level analysis of Section 6. Forty-two countries have individual-level data in Wave 7 of the World Values Survey (2017–2023) with the variables required to construct the neighborhood trust premium and the educational outcomes non-missing, and a country-level ethnic segregation index in Alesina and Zhuravskaya (2011). Countries marked with an asterisk are classified as *high ethnic segregation* (i.e., their score on that ethnic segregation index lies at or above the cross-country mean). Seven of the forty-two, marked with a dagger, do not enter the estimates of Table 4: the World Values Survey does not

Table F.2: Localization of social ties and tertiary attainment across European regions (NUTS2)

	% tertiary (25–64)		
	(1)	(2)	(3)
Localization of social ties (within-NUTS2 share)	-26.33***	-30.85***	-50.92***
	(2.392)	(4.264)	(6.315)
Distant social ties (international share)		-14.61	-12.28
		(9.792)	(8.615)
Population			✓
Observations	281	281	281
R^2	0.753	0.755	0.792
Within R^2	0.389	0.394	0.485
Country fixed effects	✓	✓	✓

Notes: NUTS2-region regressions with country fixed effects; HC1 robust SE. Outcome: share of the population aged 25–64 with tertiary education (ISCED 5–8), Eurostat 2023. Localization of social ties = within-NUTS2 share of Facebook connectedness; distant ties = share to other countries. Because this measure is defined by the region's own boundary, part of it reflects how large and how compact the region is, since a bigger or rounder NUTS2 unit retains more of its links mechanically. Country fixed effects absorb differences in how each country draws its regions but not variation in extent within a country, and the table includes no area control. The fixed-radius measures of Table F.1 and Table F.4, which count links within 50 and 25 kilometres rather than within the unit, do not have that property and give the same negative relationship with log population and log area held fixed. Developed-economy counterpart to the global region-level analysis, which lacks schooling data for high-income countries. Standard errors in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table F.3: Localization of social ties and schooling across Mexican municipalities

	Mean yrs. schooling				Attendance 15–17			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization of social ties (within-muni share)	-3.542***	-3.373***	-1.487***	-1.421***	-0.1003***	-0.1282***	-0.1255***	-0.0949***
	(0.1508)	(0.1875)	(0.1977)	(0.2038)	(0.0130)	(0.0163)	(0.0222)	(0.0221)
Distant social ties (international share)		1.807*	-2.704***	-3.341***		-0.2979***	-0.2066**	-0.2528***
		(0.9858)	(0.8620)	(0.8395)		(0.0800)	(0.0873)	(0.0875)
SES segregation (block)				0.3760***				0.0261***
				(0.0602)				(0.0055)
Localization of social ties × SES segregation (block)				-0.6209***				1.57×10^{-5}
				(0.1221)				(0.0121)
Population			✓	✓			✓	✓
Area			✓	✓			✓	✓
Observations	2,207	2,207	2,207	2,191	2,207	2,207	2,207	2,191
R^2	0.472	0.473	0.569	0.582	0.180	0.184	0.194	0.227
Within R^2	0.250	0.252	0.388	0.406	0.024	0.029	0.041	0.079
State fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Municipality-level regressions (Mexico) with state fixed effects; HC1 robust SE. Localization of social ties = share of a municipality's Facebook connectedness that stays within the municipality, so like the NUTS2 measure of Table F.2 it depends on the size and shape of the unit; distant ties = share to other countries. Scope condition (col. iv/viii): interaction with municipal SES segregation (population-weighted SD of mean years of schooling across the municipality's urban city blocks / manzanas, Census 2020). Outcomes: mean years of schooling (GRAPROES) and school attendance ages 15-17. Standard errors in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table F.4: Localization of social ties and schooling across U.S. ZIP-code areas (ZCTA)

	Mean yrs. schooling				Share BA+			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization of social ties (share ≤25 km)	-3.539***	-1.751***	-2.500***	-2.376***	-0.5356***	-0.2535***	-0.3436***	-0.3185***
	(0.0819)	(0.1104)	(0.1254)	(0.1234)	(0.0115)	(0.0159)	(0.0184)	(0.0179)
	[0.1864]	[0.2202]	[0.2360]	[0.2153]	[0.0271]	[0.0333]	[0.0369]	[0.0322]
Distant social ties (share ≥500 km)		4.055***	4.306***	4.447***		0.6400***	0.6259***	0.6530***
		(0.2122)	(0.2428)	(0.2431)		(0.0325)	(0.0370)	(0.0369)
		[0.4248]	[0.4582]	[0.4493]		[0.0627]	[0.0665]	[0.0646]
Localization of social ties × Racial segregation				-0.7887***				-0.1653***
				(0.0758)				(0.0114)
				[0.1490]				[0.0246]
Population			✓	✓			✓	✓
Area			✓	✓			✓	✓
Observations	19,561	19,561	19,561	19,433	19,561	19,561	19,561	19,433
R^2	0.499	0.520	0.533	0.534	0.584	0.607	0.612	0.617
Within R^2	0.115	0.152	0.175	0.182	0.133	0.181	0.192	0.206
County fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Notes: ZCTA-level regressions with *county* fixed effects. HC1 robust standard errors appear in parentheses and, below them in brackets, standard errors clustered on county. The segregation moderator varies only across counties, so the clustered errors are the relevant ones for the interaction; they run about twice the robust errors and leave every coefficient in the table significant at the 1 percent level. Localization of social ties = share of a ZCTA's Facebook friendship links within 25 km; distant ties = share 500 km or farther. Scope condition (col. iv/viii): interaction with county-level white/non-white residential dissimilarity across ZCTAs (its main effect is absorbed by county FE). Outcomes: mean years of schooling (from ACS educational attainment) and share with a bachelor's degree or more. Stars refer to the robust standard errors. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

record ethnic group in Germany, Japan, Jordan, New Zealand, or Turkey, and does not record parental education in Guatemala or the United Kingdom, so the corresponding fixed effects absorb those countries entirely. The estimation sample is therefore thirty-five countries, twelve of them classified as high ethnic segregation. Table F.6 re-estimates the specification without those two sets of fixed effects, which restores the full forty-two-country sample, and shows that the contrast between segregation regimes is unaffected.

Table F.5: Countries included in the World Values Survey Wave 7 sample

Continent	**Countries**
Africa	Ethiopia*, Kenya*, Morocco*, Zimbabwe*
Americas	Argentina, Bolivia, Brazil, Canada, Chile, Colombia*, Ecuador*, Guatemala*,†, Mexico*, Peru*, United States
Asia	Armenia, Bangladesh, China, India, Indonesia*, Jordan†, Japan†, Kazakhstan, Kyrgyzstan, Pakistan*, Philippines*, South Korea, Taiwan, Tajikistan, Turkey*,†, Uzbekistan, Vietnam
Europe	Czech Republic, Germany†, Greece, Netherlands, Russia*, Slovakia, Ukraine, United Kingdom†
Oceania	Australia, New Zealand†

Notes: The table lists the 42 countries for which the Alesina and Zhuravskaya (2011) country-level ethnic segregation index is available and the World Values Survey Wave 7 individual-level variables required for the analysis of Table 4 are non-missing. Countries marked with an asterisk (*) are classified as high ethnic segregation, i.e., their score on that ethnic segregation index lies at or above the cross-country mean. Countries marked with a dagger (†) are absorbed by the ethnic-group or parental-education fixed effects, which the World Values Survey does not record for them, and therefore do not contribute to Table 4; the estimation sample there is the remaining 35 countries, 12 of them high segregation. Table F.6 reports the estimates without those fixed effects, on all 42. Countries are assigned to continents following the United Nations M49 standard, with the exception that the Russian Federation is grouped with Europe and Taiwan is grouped with Asia to follow common practice in cross-country economics applications. Sampling weights provided by the WVS are applied throughout.

Table F.6: Neighborhood trust premium by segregation regime: sensitivity to the fixed-effect set

	Low ethnic segregation	High ethnic segregation	Observations	Countries
Panel A. Years of education				
Region×year, region×settlement size, sex×age	-0.011** (0.005)	-0.056*** (0.007)	64,147	42
+ country of birth, religion	-0.011** (0.005)	-0.054*** (0.007)	63,194	42
+ parental education	-0.009* (0.005)	-0.041*** (0.007)	53,730	40
+ ethnic group	-0.009* (0.005)	-0.054*** (0.007)	55,900	37
+ parental education and ethnic group	-0.011** (0.005)	-0.038*** (0.007)	46,952	35
Panel B. Upward educational mobility				
Region×year, region×settlement size, sex×age	-0.012* (0.006)	-0.032*** (0.008)	56,592	40
+ country of birth, religion	-0.012* (0.006)	-0.032*** (0.008)	55,826	40
+ parental education	-0.014** (0.006)	-0.035*** (0.008)	53,712	40
+ ethnic group	-0.013* (0.007)	-0.030*** (0.008)	48,906	35
+ parental education and ethnic group	-0.014** (0.006)	-0.034*** (0.008)	46,934	35
Panel C. Dropout before upper secondary				
Region×year, region×settlement size, sex×age	0.004 (0.005)	0.042*** (0.007)	64,147	42
+ country of birth, religion	0.004 (0.005)	0.041*** (0.007)	63,194	42
+ parental education	0.004 (0.005)	0.030*** (0.007)	53,730	40
+ ethnic group	0.003 (0.005)	0.040*** (0.007)	55,900	37
+ parental education and ethnic group	0.005 (0.005)	0.027*** (0.008)	46,952	35

Notes: Each row re-estimates the fully interacted specification of Table 4 under a different set of fixed effects and reports the slopes of the neighborhood trust premium in low- and high-segregation countries. The last row of each panel is the specification reported there. The interactions of out-group trust and of trust in close ties with the two segregation indicators are included throughout and omitted from the displayed rows. Occupation is excluded throughout: it is realized after schooling. The ethnic-group and parental-education fixed effects drop entire countries in which the World Values Survey does not record those variables, which is why the country count falls as they are added. All specifications are weighted by survey weights, and heteroskedasticity-robust standard errors appear in parentheses beside each coefficient. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Appendix References

Akerlof, George A and Rachel E Kranton, "Identity and Schooling: Some Lessons for the Economics of Education," *Journal of Economic Literature*, 12 2002, *40* (4), 1167–1201. (cited on p. 6).

Aldeco, Lucía, Mariana Calderón, Daniel Chiquiar, Gordon Hanson, Jorge Pérez Pérez, and Carlos Velázquez, "Local Labor Markets in Mexico: Definition, Databases, and Analysis," Working Paper, Banco de México, Mexico City 2024. (cited on pp. 7 and 19).

Alesina, Alberto and Ekaterina Zhuravskaya, "Segregation and the Quality of Government in a Cross Section of Countries," *American Economic Review*, August 2011, *101* (5), 1872–1911. (cited on p. 45).

Ananat, Elizabeth Oltmans, "The wrong side(s) of the tracks: The causal effects of racial segregation on urban poverty and inequality," *American Economic Journal: Applied Economics*, 4 2011, *3* (2), 34–66. (cited on p. 10).

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 12 2008, *103* (484), 1481–1495. (cited on pp. 16 and 26).

Anderson, Rodney D. and Tamara Spike, "Making History Count: The Guadalajara Census Project (1791–1930)," *Hispanic American Historical Review*, 5 2007, *87* (2), 327–351. (cited on pp. 19, 24, and 25).

Bagagli, Sara, "The (Express)Way to Segregation: Evidence from Chicag," Technical Report 2025. (cited on p. 10).

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong, "Social Connectedness: Measurement, Determinants, and Applications in Economic Research," *Journal of Economic Perspectives*, 2018, *32* (3), 259–280. (cited on pp. 43 and 46).

Baldomero-Quintana, Luis, L. Guillermo Woo-Mora, and Enrique De la Rosa-Ramos, "Infrastructures of race? Colonial indigenous segregation and contemporary land values," *Regional Science and Urban Economics*, 1 2025, *110*, 104065. (cited on pp. 5 and 14).

Bezin, Emeline and Fabien Moizeau, "Cultural dynamics, social mobility and urban segregation," *Journal of Urban Economics*, 5 2017, *99*, 173–187.

Bisin, Alberto and Thierry Verdier, ""Beyond the melting pot": Cultural transmission, marriage, and the evolution of ethnic and religious traits," *Quarterly Journal of Economics*, 8 2000, *115* (3), 955–988. (cited on p. 4).

_ **and** _ , "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, 4 2001, *97* (2), 298–319. (cited on pp. 4 and 40).

Borusyak, Kirill and Peter Hull, "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 11 2023, *91* (6), 2155–2185. (cited on p. 22).

Buringh, E., "Urbanisation hub: The clio-infra database on urban settlement sizes: 1500–2000.," 2018. (cited on p. 9).

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 11 2014, *82* (6), 2295–2326. (cited on pp. 11, 15, and 20).

Cattaneo, Matias D., Brigham R Frandsen, and Rocío Titiunik, "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate," *Journal of Causal Inference*, 3 2015, *3* (1), 1–24.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Manipulation testing based on density discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261. (cited on p. 12).

Cattaneo, Matias D., Rocío Titiunik, and Ruiqi Yu, "Estimation and Inference in Boundary Discontinuity Designs: Location-Based Methods," *arXiv Working Paper 2505.05670*, 5 2025. (cited on pp. 12, 16, and 17).

Chyn, Eric, Kareem Haggag, and Bryan A. Stuart, "The Effects of Racial Segregation on Intergenerational Mobility: Evidence from Historical Railroad Placement," *American Economic Journal: Applied Economics*, 4 2026, *18*, 34–60. (cited on pp. 10 and 45).

Cohn, Alain and Michel André Maréchal, "Priming in Economics," *Current Opinion in Psychology*, 2016, *12*, 17–21. (cited on p. 34).

de Arregui, Domingo Lázaro, *Descripción de la Nueva Galicia* 1946. (cited on pp. 8 and 10).

Glaeser, Edward L. and José A. Scheinkman, "Measuring Social Interactions," *Social Dynamics*, 12 2001, pp. 83–132.

Goldewijk, Kees Klein, Arthur Beusen, Jonathan Doelman, and Elke Stehfest, "Anthropogenic land use estimates for the Holocene - HYDE 3.2," *Earth System Science Data*, 12 2017, *9* (2), 927–953. (cited on p. 9).

IMEPLAN, "Plan de Ordenamiento Territorial Metropolitano del Area Metropolitana de Guadalajara," Technical Report, Instituto Metropolitano de Planeación, Guadalajara 2016. (cited on p. 8).

Jack, B. Kelsey, Kathryn McDermott, and Anja Sautmann, "Multiple price lists for willingness to pay elicitation," *Journal of Development Economics*, 11 2022, *159*, 102977. (cited on pp. 31, 32, 33, and 35).

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 7 1993, *60* (3), 531–542. (cited on pp. 13 and 30).

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2 2008, *142* (2), 698–714. (cited on p. 12).

Rubin, Donald B., "The Bayesian Bootstrap," *The Annals of Statistics*, 1 1981, *9* (1), 130–134.

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley, "Assessing "Neighborhood Effects": Social Processes and New Directions in Research," *Annual Review of Sociology*, 2002, *28*, 443–478. (cited on pp. 1 and 35).

Scheinkman, José A., "Social Interactions (Theory)," *The New Palgrave Dictionary of Economics*, 2008, pp. 6047–6051.

Smith, Michael E., "The archaeological study of neighborhoods and districts in ancient cities," *Journal of Anthropological Archaeology*, 6 2010, *29* (2), 137–154. (cited on p. 35).

Stephen H. Haber, "The Industrialization of Mexico, 1890-1940: The Structure and Growth of Manufacturing in an Underdeveloped Economy," *The Journal of Economic History*, 1987, *47* (2), 493–495.

Valenzuela-Casasempere, Pablo, "Displacement and Infrastructure Provision: Evidence from the Interstate Highway System," Technical Report 2025. (cited on p. 10).