

On the Other Side of the Creek:

Historical Exclusion, Neighborhood Ties, and Human Capital Gaps*

L. Guillermo Woo-Mora[†]

Job Market Paper

Version: October 9, 2026

[Click here for the most recent version.](#)

Abstract

The neighborhood ties that help households cope with historical exclusion may also raise the opportunity cost of investing in their children's education. I study this possibility in Guadalajara, Mexico, where colonial authorities settled the conquered east of a creek and the conquerors and their Indigenous allies to the west, under common municipal institutions and similar geographic conditions. Using census microdata from 1930 to 2020 and a boundary discontinuity design, I document an educational gap that has narrowed but not closed. Since 2000, school attendance has remained 3–6 percentage points lower on the historically excluded side, despite comparable school supply and quality. The gap persists among children with similarly educated parents and is concentrated in adolescence, around the transition to high school. In a pre-registered incentivized experiment, parents choose between tutoring for their children and gift cards of increasing value. Randomly asking them to draw their neighborhood beforehand reduces the share of tutoring choices by 12 percentage points (21 percent) and willingness to pay by MXN 75 (42 percent) on the excluded side, with a significantly different response across the boundary. East-side parents report similar educational aspirations and perceived returns but rely more on neighbors. Naming patterns from 1821 suggest a stronger orientation toward the local parish on the excluded side. Historical exclusion may persist not only through the resources it denied, but through the social relationships that continue to shape educational investment.

Keywords: Historical Exclusion, Persistence, Social Ties, Educational Investment.

JEL: C93, I24, N36, O15, R23, Z13.

*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, Oded Galor, 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-Cacedo, 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

Historical exclusion can outlive the rules that imposed it. To be excluded is to be denied the access and standing that others in the same society enjoy (Sen, 2000). Studies of how history shapes present-day outcomes trace the legacies of the past to institutions and public goods (Huillery, 2009; Dell, 2010), the intergenerational transmission of human capital (Wantchekon et al., 2015; Valencia-Cacedo, 2019), and residential sorting (Aaronson et al., 2021; Heblich et al., 2021). Even when access to public goods converges, the loss of standing may continue to shape whom households trust and rely on.

Excluded groups come to distrust the institutions that excluded them (Alsan and Wanamaker, 2018; Lowes and Montero, 2021; Ramos-Toro, 2023), and communities can become substitutes for the institutions their members distrust (Bowles and Gintis, 2004; Munshi, 2014). Where exclusion was drawn on a map, these communities can form around neighborhoods. In marginalized urban neighborhoods, neighbors look after each other's children, lend money, and help in emergencies (Lomnitz, 1977; González de la Rocha, 1999). This exchange is reciprocal: households receive help in return for being available to give it (González de la Rocha, 1994, 1999).

Whether neighborhood ties support or compete with educational investment is an open question. Community ties can support schooling by enforcing norms and sharing information (Coleman, 1988), but networks that reward continued participation can also discourage investments that draw members away (Portes, 1998; Munshi and Rosenzweig, 2006, 2016). Schooling is one such investment. A child who stays in school can help the neighbors less and is more likely to move away one day. Kranton (1996) shows why this kind of exchange can outlast the conditions that created it. When most people rely on favors, the market stays thin and favors remain valuable, even though everyone would be better off using the market. So the ties that protect households against exclusion can also make schooling more expensive.

This paper asks whether historical exclusion leaves a social legacy that shapes investment in human capital. I address two questions. First, when places converge in what they provide, can historical exclusion keep dividing human capital? Second, does the neighborhood affect educational investment differently where historical exclusion took place?

Answering these questions is challenging since excluded groups rarely differ from others only in standing; they also differ in ethnicity, geography, institutions, and the schools and services they receive. Because these factors move together, it is difficult to tell whether a persistent gap reflects the material or social consequences of exclusion. Households also choose where to live and help shape their neighborhoods' social ties. Differences in schooling may therefore reflect either these ties or the families who selected into the neighborhood.

I study these questions in Guadalajara, Mexico, where a decision made when the city was founded created a sharp difference in colonial standing across a creek, under common municipal institutions and similar locational fundamentals. The boundary separated the conquered, settled

in Analco, not only from the Spanish city but also from Mexicaltzingo, where the conquerors' Indigenous allies were settled, so that the comparison does not rest only on the distinction between Spanish and Indigenous populations. The material and social dimensions of exclusion have since taken different paths: no school appeared east of the creek until 1908 and school provision has since converged, but the east still carries a stigma (Jaramillo-Molina and Saucedo, 2016; Jaramillo-Molina, 2022). This allows me to ask what remains of exclusion once access to schools has converged.

I combine two empirical analyses. First, a Boundary Discontinuity (BD) design along the former creek (Cattaneo et al., 2026a), with census microdata from 1930 to 2020, documents the persistent educational gap and evaluates explanations for it. Second, a pre-registered incentivized experiment randomizes the salience of the neighborhood before parents make an educational choice, holding the household and its residence fixed, so that its effect on each side cannot reflect who lives there.

Nearly five centuries later, children on the historically excluded side are still less likely to attend school. Using the BD design, I estimate the total long-run effect of being on the east side, which includes differences in colonial status and past public provision, intergenerational transmission, and sorting, without separating them. The attendance gap was 10.6 percentage points (0.22 standard deviations) in 1930, when public provision still differed across the sides, and has stayed between 3 and 6 points (0.07 to 0.14 standard deviations) since 2000, with no detectable change as school supply and quality converged. The gap is concentrated in adolescence and is small at primary-school ages, where attendance is near universal. Over the school years, it adds up to 0.8 fewer years of schooling among adults (0.14 to 0.17 standard deviations), at the margins of completing secondary school and pursuing further education. East-side adults also earn less, by as much as their schooling predicts, and the returns to schooling are the same on both sides.

Estimating the gap at each point along the boundary (Cattaneo et al., 2026a) rules out two alternative explanations. If the avenue itself divided the city, the gap would appear along its whole length, but it is small where the boundary was urbanized only in the twentieth century and largest in the historical core. If the gap reflected the divide between Spanish and Indigenous populations, it would vanish where Analco faced Mexicaltzingo, the settlement of the conquerors' Indigenous allies, but it is about as large there as where Analco faced the Spanish city.

The usual channels of persistence account for part of this gap, but not all of it. School supply and quality, crime, and urban form show no discontinuity at the boundary, and street infrastructure is, if anything, better on the east side. The schooling gap also remains when I control for land values, so it is not simply that cheaper housing on the east side draws less-educated families. Families explain only part of the gap (Becker and Tomes, 1986; Cunha and Heckman, 2007): among children with similarly educated parents, it is still 3.7 percentage points in 2020. Household income and government transfers do not explain the remainder (Galor and Zeira, 1993; Galor, 2011), although I cannot rule out differences in wealth.

Ethnographic and survey evidence suggests that neighborhood ties may operate as an additional channel. In marginalized neighborhoods of Mexican cities, children start contributing to their households in adolescence, often at the cost of their schooling (Lomnitz, 1977; González de la Rocha, 1994). In Analco itself, residents interviewed after the 1992 explosions had lived there for thirty years on average and met their daily needs through a local economy of workshops and small shops (Reguillo, 1998). In Guadalajara, a 2022 municipality-representative survey finds that residents just east of the boundary report greater trust in neighbors and lower trust in formal institutions than residents just west of it. Residents who trust their neighbors more also have fewer years of schooling, but only on the historically excluded side.

These correlations and the ethnographic evidence do not show that neighborhood ties lower educational investment, since families who invest less may simply sort to the east side or differ in unobserved ways that shape their demand for education. Ideally, one would measure revealed demand for an educational investment while randomly varying whether parents think about their neighborhood, without changing where they live. If schooling reduces the help a household gets from its neighbors, and parents do not always account for this when they choose, making the neighborhood salient should lower their demand for education.

I test this prediction by randomizing whether parents think about their neighborhood just before an incentivized educational choice. The pre-registered experiment includes 239 parents living within 3 km of the boundary, 193 on the east side and 46 on the west, an allocation pre-specified to maximize power on the historically excluded side. Parents make 11 choices between an hour of math tutoring for their child and gift cards of increasing value, which elicits their willingness to pay for tutoring. Treated parents draw their neighborhood on a map (McCartan et al., 2024) just before these choices and control parents just after, so the map brings the neighborhood to mind without changing the household's options or its information.

On the historically excluded side, drawing the neighborhood first lowers the demand for tutoring. Parents choose tutoring 12 percentage points less often (21 percent) and are willing to pay MXN 75 less for it (42 percent). On the west side the estimates are imprecise, but the difference between the two sides is significant. When the map comes after the choice, demand is about the same on both sides. What differentiates them is how parents react when the neighborhood is on their mind. The map may make tutoring less attractive or the gift card more attractive, for instance if the card helps a family meet what it owes its neighbors. Both would make education more costly.

On the historically excluded side, the evidence suggests that neighborhood ties compete with schooling. East-side respondents get help from their neighbors twice as often as west-side respondents, and they spend more time with them. Among east-side parents, those who receive more help demand less tutoring. None of these measures changes how parents respond to the map, so I cannot pin the experimental effect on exchange with neighbors alone. A second randomized prime, the expression *de la Calzada para allá*, lowers by 15 percentage points the share of east-

side residents who explain school dropout by money or the need to work, and changes nothing on the west side.

East-side parents do not value education less. Their aspirations for their children, their expected returns to university, and their views on whether children should work or study are similar to those of west-side parents. The difference is in the bar they set: they need an 18 percent higher salary at age 30 to consider university worth it. Under simple assumptions, that premium implies a schooling gap of about 0.75 years, close to the 0.8 years in the census. The lower demand also comes from the households most involved with their neighbors, not from the poorest, which is not what a culture of poverty (Lewis, 1966) would predict.

One economic logic ties the two sets of results together. Exchange with neighbors is worth more when outside options are poor (Kranton, 1996), and school pulls children away from it (Munshi and Rosenzweig, 2006). Households that rely on their neighbors therefore invest less in education, above all in adolescents, and they respond when the neighborhood is brought to mind. Over generations, parents pass this participation on to their children (Bisin and Verdier, 2000, 2001). Exchange becomes more valuable as more neighbors participate, allowing low educational investment to persist after school provision converges, through slow adjustment or self-sustaining participation. The gap narrowed as provision and outside options improved and is largest when children begin contributing to their households. The experiment identifies a neighborhood-related response today, while the framework and historical evidence link it to long-run persistence.

Two more pieces of evidence are descriptive. In the 1821 census, men in Analco were more likely to carry the name of their parish's patron saint than men in the Spanish west or in Mexicaltzingo, who did not differ from each other. Social life on the excluded side already seems to have revolved around the local parish. After the 1992 explosions in Analco, the rebuilt blocks had adults with almost one more year of schooling than nearby blocks in 2000, and children about 5 percentage points more likely to attend school than nearby children with similarly educated parents, but both differences were gone by 2010. New residents can change a place's schooling for a while without closing its gap. Since adults' completed schooling cannot fall within a decade, the change points to a new turnover of residents; the census cannot tell whether children's attendance also fell because families adapted.

The main contribution of this paper is to show that historical exclusion can leave social arrangements that raise the cost of investing in education. Within one city, an educational gap created by exclusion outlasts the convergence of school provision and is only partly explained by families, and on the historically excluded side, bringing the neighborhood to mind lowers the demand for education while holding the household and its residence fixed. The findings suggest that equalizing what places provide need not close educational gaps when the social arrangements that exclusion left behind continue to shape the opportunity cost of educational investment.

Related literature. This paper contributes to work on long-run historical development and persistence (Nunn, 2014, 2020; Voth, 2021). Long-run differences across countries and regions have been traced to institutions (Acemoglu et al., 2001; Dell, 2010), geography (Davis and Weinstein, 2002; Bleakley and Lin, 2012), and human capital, often created by colonial investments in schooling that persisted through further public investment, transmission within families, and spillovers (Huillery, 2009; Wantchekon et al., 2015; Valencia-Caicedo, 2019). Culture is another channel: beliefs, norms, and trust shaped by historical events are transmitted across generations (Nunn, 2012; Giuliano and Nunn, 2021; Michalopoulos and Xue, 2021; Lowes, 2023). Exclusion and mistreatment in particular can leave long-run differences in trust and social preferences (Nunn and Wantchekon, 2011; Grosfeld and Zhuravskaya, 2015; Alsan and Wanamaker, 2018; Lowes and Montero, 2021; Ramos-Toro, 2023). Historical shocks also shape the layout of cities (Ambrus et al., 2020; Hanlon and Heblich, 2022; Lin and Rauch, 2022; Baldomero-Quintana et al., 2025), where historical disamenities and exclusionary boundaries persist mainly through the sorting of residents, as in the formerly polluted east sides of English cities (Heblich et al., 2021) and in neighborhoods redlined in the 1930s (Aaronson et al., 2021). I study persistence at the scale of city blocks, in the human capital of individuals rather than in house values or neighborhood status, among groups defined by where they were settled, after school provision has converged and under similar locational fundamentals. Rather than tracing the consequences of exclusion through inherited characteristics or the material organization of places, I study how the social relationships of an excluded place enter a contemporary economic decision, in an experiment that holds residence fixed.

The paper also contributes to work on how social ties shape economic decisions. In education, ties can support schooling through norms and information (Coleman, 1988), while identity, social image, and classmates shape how students invest (Akerlof and Kranton, 2002; Austen-Smith and Fryer, 2005; Sacerdote, 2011; Bursztyn and Jensen, 2015; Bursztyn et al., 2019). Beyond education, networks that provide insurance and jobs can restrict mobility (Kranton, 1996; Banerjee and Newman, 1998; Portes, 1998; Munshi, 2014; Munshi and Rosenzweig, 2016), and recent experiments measure these costs in how people earn, save, and spend (Jakiela and Ozier, 2016; Goldberg, 2017; Boltz et al., 2019; Squires, 2024; Carranza et al., 2025; Goraya et al., 2025). There is little evidence on whether the costs of belonging to such networks extend to investment in human capital. Closest to this paper, Munshi and Rosenzweig (2006) show that caste networks in Bombay, India, kept boys in schooling for traditional occupations even as returns to other schooling rose. I provide experimental evidence on this margin, where the cost falls on children and can carry gaps across generations. Work on identity and peer pressure in education looks at students and the classmates who watch them (Austen-Smith and Fryer, 2005; Bursztyn and Jensen, 2015, 2017; Bursztyn et al., 2019). I look at parents and the neighbors they trade help with, in ties that come from place, not from caste or kinship. The experiment does not separate exchange from identity or social image (Akerlof and Kranton, 2000, 2002; Shayo, 2020; Oh, 2023), but any of them would make the neighborhood a cost of educational investment.

Finally, the paper relates to work on neighborhood effects (Durlauf, 2004; Topa and Zenou, 2015; Chyn and Katz, 2021; Katz, 2026). A large literature, mostly from high-income countries, shows that where children grow up shapes their adult outcomes (Kling et al., 2007; Ludwig et al., 2013; Chetty et al., 2016; Bergman et al., 2024), with recent work emphasizing the social environment (Sampson et al., 2002; Chetty et al., 2026). In developing countries, housing lotteries and slum clearance that move poor urban households can sever informal insurance and other local ties, although networks may re-form when neighbors move together (Barnhardt et al., 2017; Franklin, 2020; Carrillo et al., 2023; Franklin, 2025; Belchior et al., 2026; Bryan et al., 2026; Rojas-Ampuero and Carrera, 2026). These studies ask what households lose when they move or leave their networks. I keep families where they are and, just before an educational choice, ask some of them to draw their subjective neighborhood (Aliprantis, 2023; McCartan et al., 2024). This does not identify the effect of exposure to a neighborhood, but it reveals a margin through which the neighborhood enters decisions, and one that differs across a historical boundary of exclusion. The 1992 explosions resemble the displacement shocks studied in this literature (Hornbeck and Naidu, 2014; Chyn, 2018; Nakamura et al., 2022), but because the census follows places rather than households, I study how the rebuilt blocks evolved rather than what happened to displaced families.

The rest of the paper proceeds as follows. Section 2 describes the context and the historical natural experiment, and Section 3 documents the persistent gap and evaluates its channels. Section 4 reports the experiment and the evidence on neighborhood ties, Section 5 develops the framework, Section 6 examines residential turnover after the 1992 explosions, and Section 7 concludes.

2 Context and Historical Natural Experiment

2.1 Guadalajara and the Colonial Boundary

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 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.² In Guadalajara, human capital varies sharply across census blocks (Figure 1b), a pattern that extends across

¹In 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 attendance beyond primary has remained above the national rate for decades. Figure A.1 in Appendix A locates Guadalajara in the national distribution, and Figure A.2 shows attendance trends since 1970.

²Among Mexicans aged 12–17, municipalities account for only 3.7 percent of the variance in attendance in 2020, while households within municipalities account for 78.7 percent. Age, sex, parental schooling, birth order, and sibship size explain 13.7 percent of the total variance. Because the harmonized census records no geography below the municipality, the household component also contains neighborhood differences. Figure A.3 reports the decomposition; Figures A.4 to A.7 extend it to parental resources, schools, and adult schooling.

the metropolitan area (Figure B.1), although residents on both sides of the historical boundary share one labor market (Aldeco et al., 2024) and face no school catchment areas.

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, implemented by 1543. Unlike central Mexico, where colonial rule co-opted the centralized Mexica (*Azteca*) state (Knight, 2002; Arias and Dirod, 2014), New Galicia was sparsely populated and politically fragmented, conquest proceeded group by group, and three earlier attempts to found Guadalajara failed amid Indigenous resistance and water shortages.³

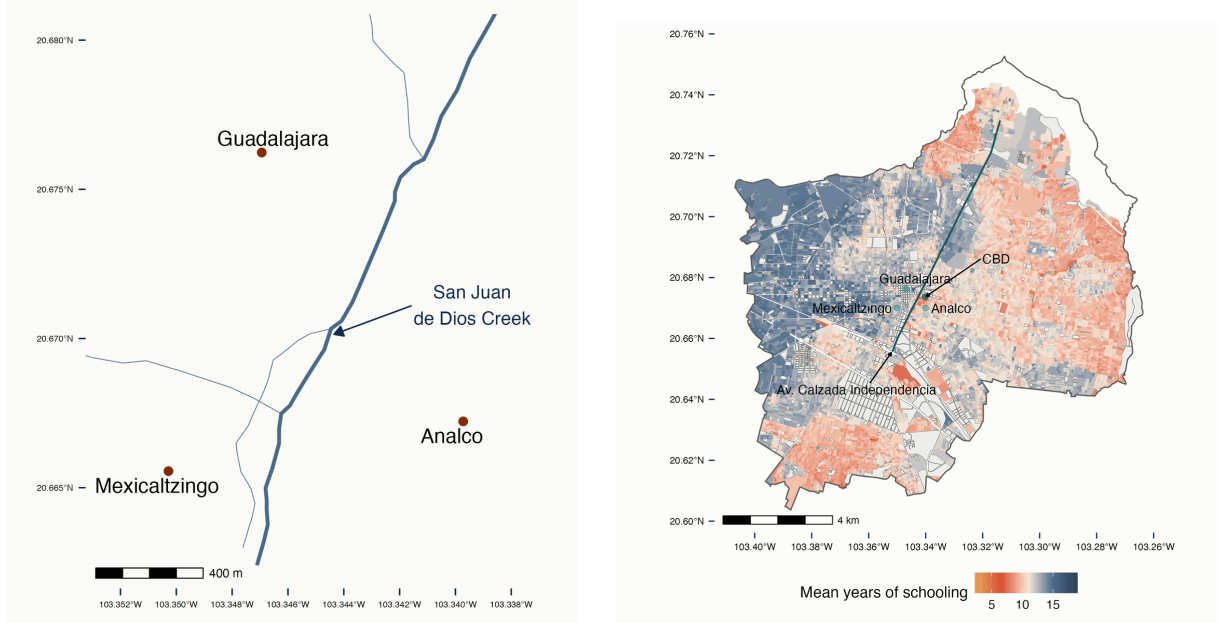
The foundation drew on a Crown-wide architecture of segregation. Beginning in 1538, the Crown established *Repúblicas de Indios*, corporate organizations for Indigenous populations with their own *cabildos* and communal land rights, distinct from *Pueblos de Indios*, the settlements where Indigenous populations lived (Castro Gutiérrez, 2010; Graham, 2013). The system sought spatial separation between Indigenous and Spanish populations, although enforcement in Guadalajara was partial (Anderson, 1988), and it was abolished after independence in 1821.

Within Guadalajara, this 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 south on the same bank in *Mexicaltzingo* ("house of the Mexicas"). In 1543, Franciscans established *Analco* ("on the other side of the creek" in Nahuatl) east of the creek as a *Pueblo de Indios* for the conquered local population (Solís Matías, 1986, 1992). Figure 1a maps the three settlements, all within roughly one kilometer of each other. The two Indigenous settlements were not symmetric, and that distinction is central to interpreting the design. Mexicaltzingo's settlers had participated in the conquest and were recognized as allies, receiving tribute exemptions, communal land grants, and recognition of their leaders as *principales*. They were "*not conquered, but among the conquerors*" (Castro Gutiérrez, 2010). Analco, the only foundational settlement east of the creek, was populated by conquered local groups and later regional migrants. The east–west contrast is therefore one of colonial status, with conquerors and their Indigenous allies on the west and conquered groups on the east, rather than a Spanish-versus-Indigenous divide.⁴

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

⁴Reading the contrast through modern ethnic categories would be anachronistic. Alberro (2019) argues that colonial New Spain classified people through categories such as estate, *nación*, and corporate membership rather than race, and that *indio* was itself a porous category. The sources distinguish Mexicaltzingo and Analco by standing within that corporate order, not simply ethnic origin.

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 because 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 suppressed by INEGI because of small population counts.

A historical natural experiment of neighborhood formation. Four features of the foundation, together with the later transformation of the boundary in Section 2.2, support its interpretation as a historical natural experiment in the sense of Cantoni and Yuchtman (2021); Appendix A reports the supporting evidence.

First, settlement locations were unrelated to pre-colonial agglomerations, which can shape economic activity over long horizons (Davis and Weinstein, 2002; Bleakley and Lin, 2012; Maloney and Valencia-Caicedo, 2016; Lee and Lin, 2018). The region had no urban centers before Spanish contact, and pre-colonial population estimates (Goldewijk et al., 2017) and colonial-era urban density data (Buringh, 2018) show sparse population relative to central Mexico through the 16th–18th centuries (Figures A.8 and A.9).

Second, placement was not driven by geographic fundamentals. Altitude, slope, and climate are similar across banks, and a BD design within 1.5 km of the SJdD creek finds no significant discontinuities in locational characteristics (Figure A.10 and Table A.1).

Third, the settlements shared the same political and municipal institutions from the start (Acemoglu et al., 2001, 2005; Dell, 2010; Lowes et al., 2017). All fell under colonial authorities centered in Guadalajara and the Audiencia of New Galicia (Lázaro de Arregui, 1946; Solís Matías, 1992), and neither Indigenous settlement held the corporate autonomy of a *República de Indios*:

both were governed by the city's *ayuntamiento*.⁵ The settlements differed in status and privileges within these institutions, not in the institutions themselves, and they have remained within the same jurisdiction ever since.

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

Together, these features make the colonial assignment a credible source of variation in exclusion within a common physical, municipal, and market environment.

2.2 Historical Exclusion and Its Two Dimensions

Exclusion. The colonial assignment is a case of historical exclusion as defined in the introduction. Here, exclusion took an administrative form: an authority assigned a group a status that denied it the standing granted to comparable groups within the same jurisdiction (Sen, 2000; Ramos-Toro, 2023). I examine two dimensions of that exclusion: a material one, public provision, and a social one, the standing of each side and the organization of local social and economic life.

The material dimension. The material dimension of exclusion is visible in the location of the city's institutions. Four general maps of Guadalajara, from 1800, 1863, 1890, and 1908, name institutions and locate them on the urban grid (Figure A.12). The eastern bank contained between a third and two-fifths of the population but less than fifteen percent of named institutions on every map, and no educational institution appears east of the creek until a girls' school run by a religious order in 1908 (Figure A.11 and Table A.2 in Appendix A.4). Because the maps record institutions their authors considered worth naming, they describe the location of named institutions rather than the complete supply of schooling or other public goods. Assessed land values in the 1882 cadaster were also roughly one-half to two-thirds lower on the eastern bank at the creek (Figure A.14). Educational provision has since converged: today, school supply and quality show no discontinuity at the boundary (Section 3).

From the creek to the Calzada. The boundary is not a physical barrier today either. The SJdD creek was piped underground between 1903 and 1911 and replaced by Calzada Independencia, a surface-level avenue that, unlike the railroad tracks and urban freeways studied as infrastructural barriers (Ananat, 2011; Chyn et al., 2026; Bagagli, 2025; Valenzuela-Casasempere, 2025), connects the two banks rather than separating them (Figure A.16 in Appendix A.6). A discontinuity at the boundary is therefore unlikely to reflect the cost of crossing it.

⁵Around 1800, Analco and Mexicaltzingo are two of four *pueblos de indios* near Guadalajara absent from the *propios y arbitrios* accounts submitted to the Contaduría. Tanck de Estrada (1999) interprets this absence as evidence that they fell under the city's *ayuntamiento*.

The social dimension. The social dimension of exclusion, by contrast, remains in two forms. First, the difference in standing has persisted. By 1821, both banks were ethnically mixed but still differed in rank (Figure A.18 and Table A.3). The early twentieth-century press associated the eastern side with disrepute (Trujillo Bretón, 2018, pp. 310–311), and residents today locate themselves and others with the expressions *de la Calzada para allá* and *del otro lado de la Calzada*, both referring to the eastern bank and marking it with a stigma (Jaramillo-Molina and Saucedo, 2016; Jaramillo-Molina, 2022); Reguillo (1998) likewise describes the piped creek as a symbolic mark dividing the city. Second, local social and economic life on the eastern bank has been organized around neighbors (Section 4). Reguillo’s ethnography of Analco reports residents who located risk outside the neighborhood, organized their local economy around relations with neighbors, and had lived there for thirty years on average.

3 Persistent Human Capital Gaps

Nearly five centuries later, the historical shock predicts a local divergence in human capital at a fine spatial scale, even after the school supply and school quality have largely converged on both sides of the boundary. The boundary divides not only the stock of schooling, but also the decision to attend school.

3.1 Boundary Discontinuity Design and Identification

Geolocated census microdata. For the main analysis, I use individual-level, geolocated census microdata from the Mexican national statistical agency, INEGI.⁶ I exploit the 2020 *Censo de Población y Vivienda*, which includes the universe of individuals residing in Guadalajara and their location of residence down to the census block, allowing to run the BD design at an extremely local scale within meters of the historical boundary. The census collects information on parental education and several individual, family, and household variables, but does not include income. I complement the analysis with data from the 2015 *Encuesta Intercensal*, a large survey representative at the municipality level, that includes income and similar parental and household variables as the census, but has few observations near the boundary, so I use it only for municipality-wide comparisons.

Specification. I implement a BD design (Cattaneo et al., 2026a) around the SJdD creek, today Calzada Independencia.⁷ The baseline specification collapses location into the signed orthogonal

⁶These 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.

⁷The BD design is the regression discontinuity design that applied work often calls a spatial or geographic RD design: units are assigned to treatment by the location of their residence relative to a boundary, so the score has two dimensions instead of one. Cattaneo et al. (2026a) show that pooling observations by their distance to the boundary

distance to the boundary:

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

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, and the running variable s_i is the orthogonal distance in meters to the boundary, positive on the eastern side. The interaction lets the outcome gradient differ across banks. The control vector X_i includes sex, age, longitude, latitude, and boundary-segment fixed effects, so that comparisons occur within the same segment (Dell, 2010).

The estimation follows the recommendations of Cattaneo et al. (2026a) for BD designs. They show that interacting distance with treatment is crucial to allow the fit to differ across sides, and they recommend MSE-optimal bandwidths with robust bias-corrected inference, which I implement following Calonico et al. (2014). Because treatment changes along a line rather than at a point (Keele and Titiunik, 2015), τ estimates the boundary average treatment effect, a density-weighted average of the effects at each point of the boundary. Distance-based estimates can be biased when the boundary is irregular, as a natural creek is (Cattaneo et al., 2026b), so Cattaneo et al. (2026a) recommend location-based methods that use both coordinates. I therefore also estimate the boundary average treatment effect curve, the effect at each point of the boundary, with the bivariate estimator of Cattaneo et al. (2025), and report it in Section 3.4.

Additionally, I estimate the gap in school attendance conditional on parents' education and some household characteristics (size and migration status). Since some of these controls may have been affected by the historical assignment, these estimates are not interpreted as the total effect of the assignment, but as the gap that remains among children from comparable households.

Identification and interpretation. Identification of the effect of the historical assignment requires that, had the historical assignment not existed, the average potential outcomes would be continuous at the cutoff. This assumption is hard to come by in the presence of geographical boundaries, as people can choose where to live (e.g., Lee and Lemieux, 2010; Keele et al., 2017).

Three features of the setting rule out sources of discontinuity other than residential sorting. First, the boundary pre-dates individuals' current housing locations, as the area was not urban in 1542, and the geographical characteristics of the area (terrain, altitude, slope, climate, and hydrology) are balanced on both sides of the creek (Section 2). Second, the boundary is not a "compound treatment" in the sense of Keele and Titiunik (2015): it does not delimit administrative or school districts;⁸ both sides of the creek are part of the same municipality, fall under the

recovers a density-weighted average of effects that can vary along it, and they provide methods to estimate the effect at each point of the boundary.

⁸Schools are organized into "zonas escolares" by the Secretaría de Educación del Estado de Jalisco to streamline inspection, plan education policy, and assign new school openings. However, they do not follow a strict catchment area rule.

same court, fiscal, and police jurisdictions, and share the same labor market, and the Calzada does not represent a physical barrier to mobility between the two sides (Figure A.17). Third, present-day ethnicity does not define an alternative partition at the boundary: there is a low prevalence of Indigenous population (according to self-identification or language) in Guadalajara, and (continuous) measures of individuals' skin tone (interviewer's perception) are not clustered by side of the boundary (Figure A.19).

Residential sorting, however, cannot be ruled out. There is a discontinuity in the density of individuals' residential location at the cutoff: there are more individuals on the east side of the boundary, although the relative gap decreases to 4% at 500 meters from the cutoff (Figure B.2; McCrary, 2008; Cattaneo et al., 2018). This seems to be related to the fact that the commercial area of Guadalajara is located on the west side of the avenue, and residential use starts on the east side. Thus, I do not interpret this discontinuity as suggestive of manipulation of the running variable, but it leaves open the possibility of residential sorting.

Under the continuity assumption, I interpret τ as the reduced-form long-term impact of the historical assignment, capturing not only the impact of differential initial status and public good provision, but also intergenerational transmission, residential sorting, and any other differential social mechanisms that may have arisen alongside the historical assignment. I discuss some of these mechanisms, from supply, household, and sorting margins, in the following subsections, without making causal claims.

3.2 The Gap Today

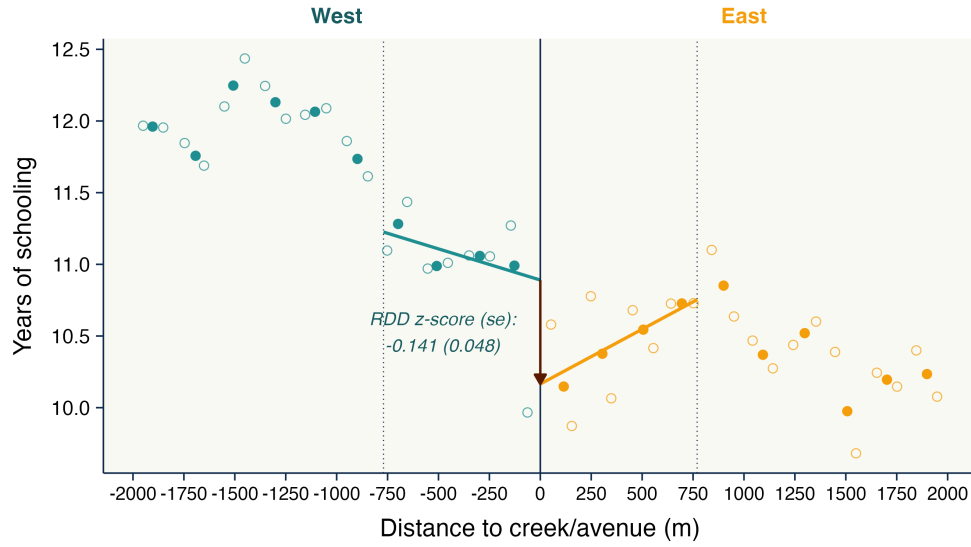
The boundary still divides both completed schooling and current attendance. Completed schooling is a stock variable, accumulated over multiple cohorts, and driven by both the history of school provision and the intergenerational transmission of schooling. Current school attendance, on the other hand, reflects the choices of current households, in a context where school provision has converged. The rest of this section investigates this outcome.

Panel (a) of Figure 2 shows that residents immediately east of Calzada Independencia attain about 0.8 fewer years of schooling than residents immediately west, 0.17 standard deviations or about 9 percent of the municipal mean. The discontinuity is visible within 100-meter bins on either side of the boundary, and it is 0.14 standard deviations with covariates. Panel (b) shows that school attendance among individuals aged 6–22, an age range chosen to ensure sufficient observations near the boundary, drops by 3.2 percentage points at the boundary ($p = 0.09$). Table B.1 reports the estimates across polynomial orders, kernels, variance estimators, and covariate sets.

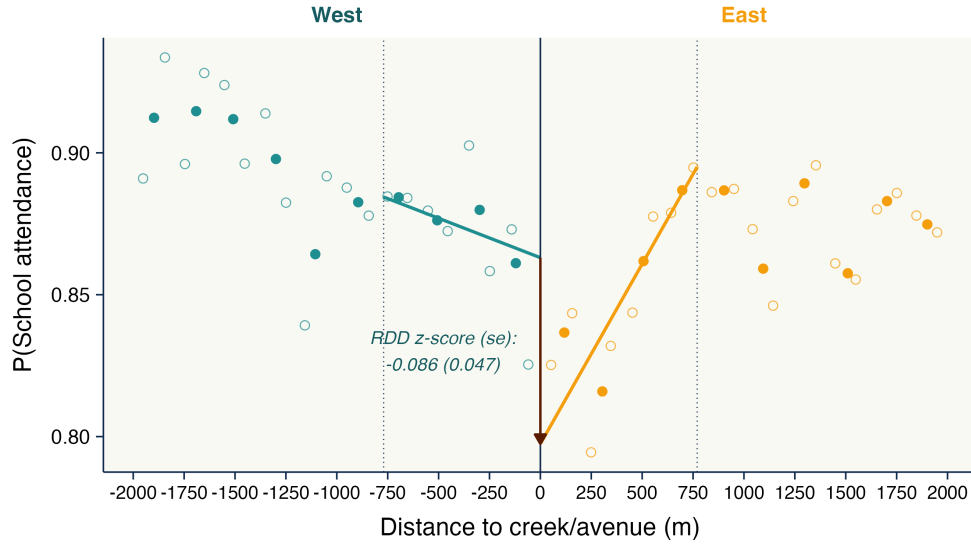
3.3 The Gap over Time

The gap was already present in 1930, narrowed during the twentieth century, and has not closed since 2000. I use the 1930 census at the *cuartel* level, the 1990 census at the AGEB level, and

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



(a) Years of schooling, ages 6+



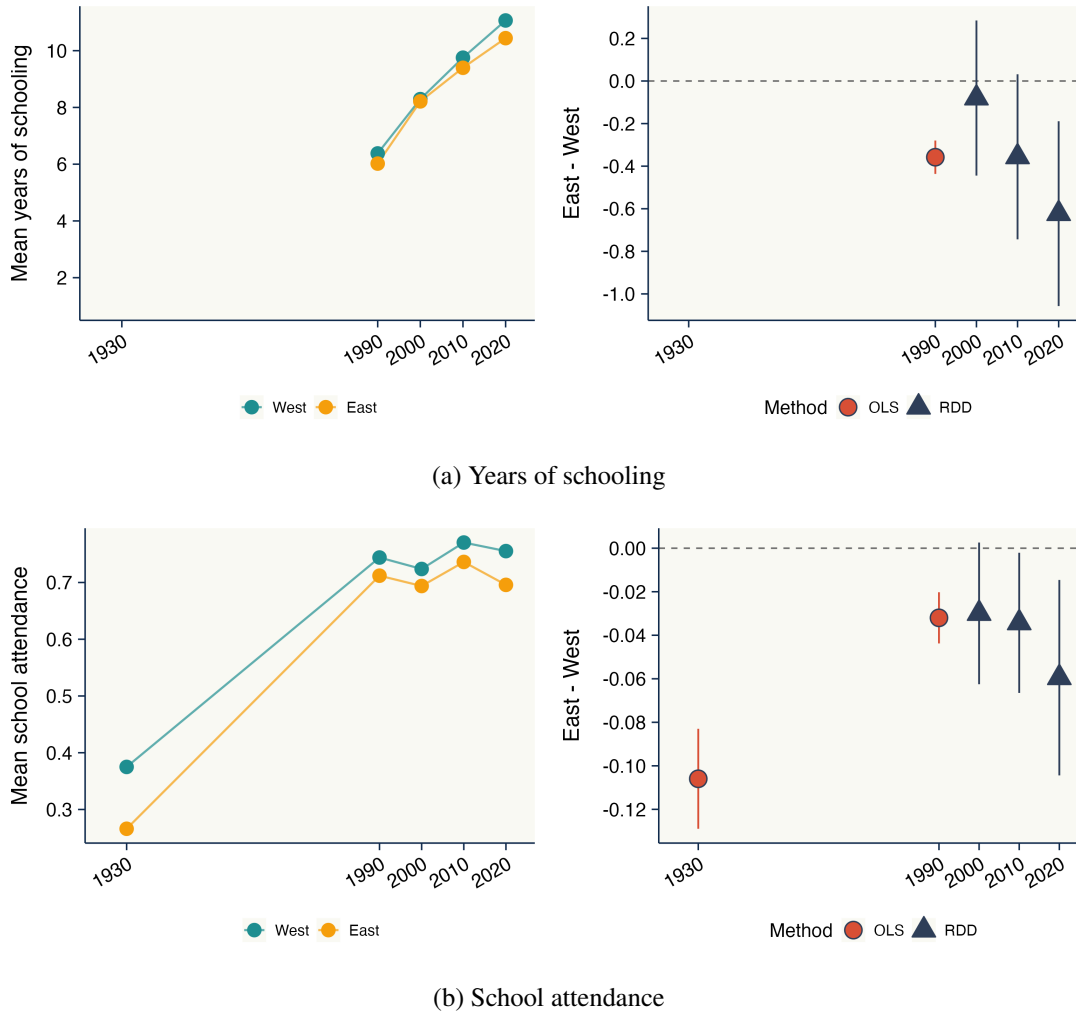
(b) School attendance, ages 6–22

Notes: The figures display binned means and local linear fits of the outcome variables using individual-level data from the complete 2020 *Censo de Población y Vivienda* for the municipality of Guadalajara. The x-axis measures the (signed) orthogonal distance in meters to the historical boundary, where 0 indicates the current location of the Calzada Independencia, a former creek (San Juan de Dios); positive values are on the historical east side of the boundary (in Analco). The y-axes measure the outcome variables in their original units: average years of schooling (Panel (a), for individuals aged 6 or above) and the probability of attending school (Panel (b), for individuals aged 6 to 22). Hollow (solid) circles correspond to the average outcome of individuals living within 100 (200) meters of the historical boundary; the x-axis binning is restricted to $[-2000, +2000]$ meters. The vertical dotted lines indicate the bias-corrected bandwidth $b = \pm 770$ meters for years of schooling chosen as in Calonico et al. (2014). The thick colored lines show local linear fits restricted to the same bandwidth on both sides of the boundary (and without covariates). The coefficients in the figures, instead, correspond to the covariate-adjusted, standardized estimates in Table B.1 with their respective MSE-optimal estimation bandwidth h (395 meters for Panel (a) and 895 meters for Panel (b)), controlling for sex, age, longitude, latitude, boundary segment fixed effects, and – in Panel (b) – parental education, household size, and migration. The local linear fits and the annotated estimates are thus not based on the same specifications. No additional curve or band is shown. The figure indicates a relative human capital disadvantage of individuals living on the Analco side of the boundary.

the 2000, 2010, and 2020 censuses at the block level.⁹

⁹For 1930 and 1990, I compare *cuarteles*/AGEBs immediately adjacent to the boundary using a parametric specification. For 2000–2020, I estimate the non-parametric BD specification of Equation (1) at the block level,

Figure 3: Human capital gaps across the historical boundary over time (1930–2020)



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 BD 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 BD 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 1990 attendance estimate also conditions on parental schooling. The series restricts each census to blocks within five kilometers of the boundary and therefore differs from the estimates in Section 3.2, which use all blocks of the municipality: in 2020, the schooling gap without covariates is –0.62 years in this series and –0.79 years in Table B.1, and the corresponding attendance gaps are 6.0 and 3.2 percentage points. Adult literacy is reported in Figure B.3.

In 1930, when public provision still differed across the sides, residents aged 18 or younger in the restricting each census to blocks within five kilometers of the boundary.

cuarteles adjacent to the boundary were 10.6 percentage points less likely to attend school on the Analco side than on the west, with age and sex fixed effects (Table B.2).¹⁰ Adult literacy was 56.8 percent on the east against 68.6 percent on the west. The educational divide was therefore already present before the major local and federal expansions of public schooling in Mexico.

Figure 3 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 recent evolution. The literacy gap closed as literacy approached universal coverage and shows no discontinuity from 1990 onward (Figure B.3). The attendance gap fell to about 3 percentage points in 2000 and 2010, and I cannot reject that it is unchanged between 2000 and 2020. The schooling gap, if anything, widens over the same period ($p = 0.06$), from -0.08 years in 2000 to -0.62 years in 2020 in this series.¹¹ The boundary therefore marked an educational divide by 1930 that narrowed as public education expanded but remains visible after school supply and quality converged.

3.4 The Gap along the Boundary

The gap is largest in the historical core, where Analco faced the Spanish city at one point of the boundary and Mexicaltzingo at the next. An ethnic reading of the east–west contrast would predict a narrower gap where Analco faced Mexicaltzingo, an Indigenous settlement whose residents arrived as allies of the conquerors. An explanation based on the avenue itself would predict a similar gap along its whole length. The northernmost stretch lay outside the nineteenth-century city and was urbanized only in the twentieth century (Figure A.15).¹²

I estimate the boundary average treatment effect curve using the location-based estimator of Cattaneo et al. (2025), with uniform confidence bands along the boundary. At points 4 and 5, where Analco faced the Spanish city (point 4) and Mexicaltzingo (point 5), the estimates are -2.3 and -1.9 years of schooling and -12.6 and -11.8 percentage points in school attendance (Table B.12). The discontinuity is larger at each of these two points than at each of the three northern ones, for years of schooling and for school attendance alike.¹³ The two points do not differ from each other, so the gap is as large facing the Spanish city as facing the conquerors’

¹⁰Analco comprises *cuarteles* 8 and 9 on the east; central Guadalajara and the historical Mexicaltzingo area comprise *cuarteles* 1, 2, and 7 on the west. The estimate depends on this choice: excluding Mexicaltzingo raises the gap to 19.1 percentage points, and including all *cuarteles* lowers it to 5.8.

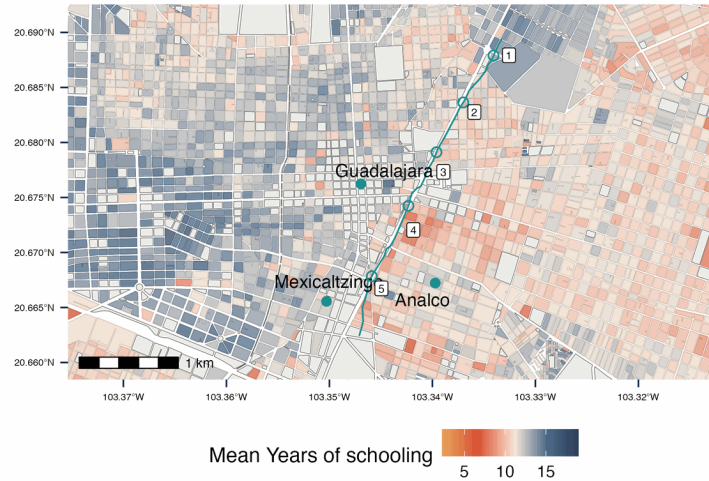
¹¹The near-zero schooling estimate in 2000 may partly reflect the 1992 Reforma explosions, which destroyed and repopulated part of Analco within the BD bandwidth shortly before the census. Section 6 shows that the rebuilt blocks were more educated than the rest of Analco in 2000 and converged afterward.

¹²The two northernmost boundary points lie more than 200 meters from the nearest block of the 1890 city map, and the other three lie within 60 meters of one.

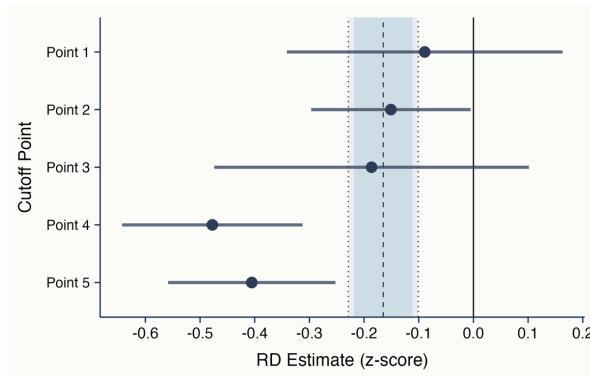
¹³Five of the six pairwise contrasts are significant at the 5 percent level for each outcome and the sixth at the 10 percent level; all sharpened false-discovery-rate q -values are below 0.03 (Anderson, 2008). Because rd2d does not report covariance across boundary points, the standard errors of the contrasts treat the estimates as independent. One additional boundary point lacks a bias-corrected estimate and is omitted.

Indigenous allies. Along the stretch urbanized in the twentieth century (points 1 and 2), the schooling estimates are -0.4 and -0.7 years and the attendance gap is indistinguishable from zero. At point 3, the estimates are small because observations are sparse, as that stretch borders Guadalajara's historic commercial core, where few people live on the western side.

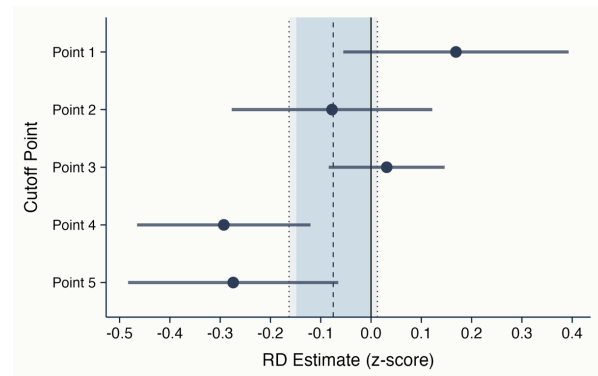
Figure 4: Boundary average treatment effect curve (2020)



(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 the boundary average treatment effect curve 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 distance-based BD estimate from Figure 2; shaded areas show the 95% and 90% confidence intervals.

This pattern is difficult to reconcile with either the avenue itself or Indigenous ancestry generating the gap, although it does not separately identify colonial status from every other historical difference between the settlements.

3.5 Standard Explanations

Standard determinants of educational investment account for part of the attendance gap, but not all of it. Parental human capital, household resources and credit constraints, residential sorting,

school quality, and local conditions can all sustain differences in educational investment (Becker and Tomes, 1986; Cunha and Heckman, 2007; Galor and Zeira, 1993; Galor, 2011). I examine each in turn.

Parental human capital. Parental human capital explains at most a third of the attendance gap. In 1930, the census records literacy but not years of schooling, and among residents whose household head's literacy is recorded, conditioning on it reduces the gap from 10.9 to 9.7 percentage points, or by 11 percent. In 2020, conditioning on parental schooling, household size, and migration status leaves an attendance discontinuity of 3.7 percentage points, against 3.2 without them (Table B.1). In the 2015 Intercensal Survey, parental schooling reduces the gap among adolescents aged 12–18 from 8.1 to 5.1 percentage points, about a third, and household income does not reduce it further (Table 1).¹⁴ Among children aged 6–11 with similarly educated parents, the gap is 1.5 percentage points, so it opens in adolescence. The Intercensal comparison separates ages 12–18 from ages 6–11, and the public block tabulations of the 2020 Census narrow the timing to the transition to high school: within three kilometers of the boundary, attendance is 3.9 percentage points lower on the east side at ages 15–17 and is not lower at ages 6–14. Because parental schooling may itself reflect the historical assignment, these comparisons are not a mediation exercise, and they do not remove other forms of family transmission, such as skills or aspirations (Cunha and Heckman, 2007; Genicot and Ray, 2020). A substantial gap remains among children whose parents have similar schooling.

Household resources and returns to schooling. Measured household resources and the returns to schooling do not explain the remaining gap. In the Intercensal Survey, adding household size and current income leaves the adolescent attendance gap at about 5 percentage points. Household income, government transfers, remittances, pensions, housing finance, and employment-linked housing credit also do not jump at the boundary (Table B.3 in Appendix B.6). Adults east of the boundary earn about 8 percent less per month, but the gap falls to 1 percent and becomes statistically insignificant after controlling for schooling, and the interaction between Analco and schooling is small and insignificant. Consistent with Guadalajara forming a single labor market (Aldeco et al., 2024), the returns to schooling are similar across banks. These measures do not capture wealth, liquidity, or historical credit constraints, so they do not rule out those channels.

Housing-market sorting. A schooling gap remains after controlling for land values, although sorting on unobserved characteristics cannot be ruled out. If housing prices capitalize local amenities or the historical stigma of the boundary, households with different resources or pref-

¹⁴Because the Intercensal sample thins quickly near the boundary, I estimate a parametric east–west gap with side-specific linear trends in distance, so the coefficient on the Analco indicator is the fitted gap at the boundary. This specification relies on households farther from the creek and is therefore more exposed than the local BD design to sorting, provision differences, and functional form.

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

	Years of schooling	$\mathbb{1}\{\textit{Attends school}_i\}$						Monthly income (log)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathbb{1}\{\textit{Analco}_i\}$	-1.31*** (0.118)	-0.052*** (0.015)	-0.081*** (0.024)	-0.054** (0.024)	-0.051** (0.023)	-0.052** (0.025)	-0.015* (0.008)	-0.084*** (0.027)	-0.010 (0.024)	0.075 (0.068)
Years of schooling (YoS)									0.065*** (0.001)	0.068*** (0.005)
$\mathbb{1}\{\textit{Analco}_i\} \times \text{YoS}$										-0.006 (0.006)
Parental YoS				0.021*** (0.001)	0.018*** (0.001)	0.019*** (0.001)	0.003*** (0.0006)			
Household income (log)						0.0002 (0.008)	0.0000 (0.002)			
Sample (age in years)	6+	6–18	12–18	12–18	12–18	12–18	6–11	18+	18+	18+
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Household size					✓	✓	✓			
Observations	74,779	17,463	9,922	9,916	9,916	9,041	8,059	31,584	31,478	31,478
Dep. var. mean	9.18	0.884	0.807	0.807	0.807	0.803	0.984	8.65	8.65	8.65
Dep. var. SD	4.71	0.320	0.395	0.395	0.395	0.398	0.127	0.644	0.645	0.645
R ²	0.443	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 $\mathbb{1}\{\textit{Analco}_i\}$ 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 six columns report the probability of attending school for the indicated age range. The fourth through seventh columns additionally control for the maximum of mother's and father's years of schooling; the fifth through seventh for household size; and the sixth and seventh 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.

erences may sort across space, as [Baldomero-Quintana et al. \(2025\)](#) document for Mexico City. Parcel-level cadastral data show a clear west–east gradient in land values (Figure B.4a), which correlate strongly with online housing listings (Figure B.4b). Re-estimating the schooling discontinuity at the block level while controlling for land values yields gaps from 0.80 to 0.97 years, close to the 0.79-year individual-level estimate (Tables B.4 and B.5 and Figure B.5).¹⁵ Because land values may themselves capitalize the boundary's history, conditioning on them may absorb part of the effect of interest.

School supply and quality. Current school supply and quality do not differ at the boundary. I find no discontinuity in schools per block, sector, level, shift, public-school share, or students per teacher (Table B.6). Administrative records show lower approval rates, higher repetition, and larger classes on the eastern bank through about 2002, but these gaps close by the mid-2000s, and from then onward neither administrative indicators nor ENLACE scores differ across the

¹⁵This exercise covers schooling only because the public block-level census suppresses blocks with few households, preventing a comparable attendance measure. Within the block-level specification, the standardized gap falls from -0.62 to between -0.40 and -0.48 standard deviations.

boundary (Figures B.6 to B.9). The 2020 attendance gap therefore concerns cohorts schooled mostly after this convergence, while earlier quality differences are one of the historical channels bundled into the schooling stock. Because schools have no strict residential catchments, measured supply at the boundary also understates the choice set available to households on both sides.

Crime exposure. Crime does not differ at the boundary. Using more than 187,000 georeferenced crime reports from 2017–2024, I find no significant discontinuity in overall crime or any major category, including homicide (Table B.7 and Figure B.10).

Street infrastructure, urban form, and businesses. Street infrastructure is not worse east of the boundary and, where block fronts differ, the eastern side has more paving, sidewalks, and ramps (Table B.8). Residents and dwellings per block and parcel size do not differ (Table B.9), and establishments east of the boundary are smaller but have the same mix of activities (Tables B.10 and B.11).

What remains. Parental human capital, measured household resources, the returns to schooling, and current school provision do not exhaust the contemporary attendance gap. This does not make them unimportant: parental education accounts for up to a third of the gap, and historical school provision contributes to the schooling stock observed today. It leaves room for an additional demand-side margin, which Section 4 examines.

3.6 Robustness

The schooling and attendance gaps survive several diagnostic checks; full results are in Appendix B.

Migration. In 2020, 91.5 percent of Guadalajara residents lived in Jalisco five years earlier, and excluding inter-state migrants yields nearly identical estimates (Table B.1). The census follows places rather than households, so it cannot bound the selection of residents who left; Appendix B.12 reports a proxy for residential persistence and its limitations.

Donut estimates. Because residential density differs at the boundary, the innermost observations come disproportionately from the eastern side. Dropping every resident within 50, 100, or 150 meters of the line leaves the schooling gap between -0.10 and -0.21 standard deviations and the attendance gap between -0.09 and -0.13 (Table B.14).

Spatial autocorrelation and placebo boundaries. Following Müller and Watson (2024), and using the package by Becker et al. (2025), I reject a spatial unit root and fail to reject spatial stationarity, supporting local BD estimators without spatial differencing. Randomly shifting

and rotating the boundary 1,000 times within the municipality (Figure B.11) yields a placebo distribution 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 The Neighborhood and the Demand for Education: Experimental Evidence

Section 3 leaves an attendance gap that parental schooling, household resources, and school provision do not explain. This section shows that the neighborhood is a margin of educational demand: making it salient lowers parents' demand for tutoring on the historically excluded side of the boundary, and not on the other.

4.1 Qualitative and Survey Evidence on the Neighborhood

Urban ethnography in Mexico documents marginalized households that rely on reciprocal exchange with neighbors and kin to get by. In a Mexico City settlement, Lomnitz (1977) found that households survived recurrent periods without income through exchange networks among neighbors, sustained by trust and physical proximity, and that families whose incomes rose stopped asking for help and eventually moved away. In working-class Guadalajara of the 1980s, household survival depended on pooling the work of several members, with children starting to work around age 15, and on support from relatives and neighbors (González de la Rocha, 1994). This support is costly to keep: households must be available when others need help and have resources to reciprocate, so the poorest are often left out (González de la Rocha, 1999).

The historically excluded side of the boundary in Guadalajara shares these patterns. In Analco, residents interviewed after the 1992 explosions had lived in the neighborhood for thirty years on average, twice as long as in the other affected areas, and a local economy of workshops and small shops that survived the relocation of the city's bus terminal let them meet daily needs without leaving the neighborhood (Reguillo, 1998). Residents of the east side also attribute the difficulty of getting ahead to their neighbors or their *barrio* (Jaramillo-Molina, 2022).

A representative survey. A representative survey of Guadalajara residents points in the same direction: residents just east of the boundary place more trust in their neighbors and less in formal institutions, and among them, trust in neighbors is associated with less schooling. I use the 2022 Life Quality Survey, which interviews 289 residents in 23 census tracts across the municipality,¹⁶ and estimate discontinuities at the boundary with a separate linear distance gradient on each side. Table 2 shows the results.

¹⁶The survey is conducted by the NGO *Jalisco Cómo Vamos*.

Residents just east of the boundary are 37 percentage points more likely to report high trust in their neighbors, but not in their family, coworkers or classmates, or people on the street, so the difference is specific to neighbors. The difference appears in every age bracket, including respondents aged 18–29 (Figure C.1). They also report lower trust in formal institutions, by 0.23 on a 0–1 index, consistent with exclusion shifting social life inward (Table C.1). Membership in neighborhood, school, political, religious, or assistance organizations does not differ, so the difference lies in informal trust toward neighbors rather than in formal participation (Table C.2). These patterns preceded the experiment and shaped its design: they suggest that, on the east side, social life is organized around neighbors in a way that may compete with educational investment.

Table 2: Trust, participation, and schooling at the boundary: Life Quality Survey (2022)

	Trust in				Institutional trust index	Member of					Years of schooling
	Neighbors	Family	Coworkers or classmates	People on the street		Neighborhood assoc.	School-related org.	Political org.	Religious group	NGO or assistance org.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
<i>Panel A. Without schooling control</i>											
$1\{Analco_i\}$	0.372** (0.149)	0.072 (0.110)	0.134 (0.158)	0.049 (0.129)	-0.235*** (0.064)	0.066 (0.056)	-0.039 (0.050)	0.023 (0.041)	-0.076 (0.085)	-0.062 (0.063)	-0.798 (1.23)
Trust in neighbors											0.796 (0.691)
Trust in neighbors $\times 1\{Analco_i\}$											-1.74* (0.945)
<i>q</i> -value	[0.035]	[0.390]	[0.329]	[0.481]	[0.004]	[0.268]	[0.347]	[0.399]	[0.329]	[0.328]	
<i>Panel B. Controlling for years of schooling</i>											
$1\{Analco_i\}$	0.368** (0.151)	0.084 (0.111)	0.165 (0.156)	0.035 (0.129)	-0.232*** (0.065)	0.080 (0.057)	-0.032 (0.050)	0.032 (0.044)	-0.064 (0.087)	-0.057 (0.063)	
<i>q</i> -value	[0.041]	[0.404]	[0.295]	[0.480]	[0.006]	[0.200]	[0.404]	[0.404]	[0.404]	[0.361]	
Observations	289	289	289	289	289	289	289	289	289	289	289
Dep. var. mean	0.655	0.912	0.585	0.217	0.229	0.041	0.018	0.036	0.065	0.018	11.33
Age–sex fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey weights	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimates from the 2022 Life Quality Survey of Jalisco *Cómo Vamos*, representative of Guadalajara residents (289 respondents in 23 census tracts). Each column regresses the outcome on an indicator for residence east of the historical boundary and separate linear distance gradients on each side, with age–sex fixed effects and survey weights. The outcomes in columns (1)–(4) are indicators for reporting high trust in each group, and those in columns (6)–(10) are indicators for membership in each type of organization. Column (11) adds trust in neighbors and its interaction with the Analco indicator. Robust standard errors in parentheses. Sharpened *q*-values (Anderson, 2008) adjust for the 26 trust and participation outcomes of the survey within each panel. The estimates extrapolate distance gradients toward the boundary from tracts at a median distance of 3.7 kilometers. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

4.2 Reciprocal Exchange and the Experimental Prediction

Reciprocal exchange with neighbors has economic benefits. Neighbors look after each other’s children, lend money and tools, share information about jobs, and step in during illness or emergencies. These exchanges are reciprocal: receiving help depends on remaining available to provide it (Kranton, 1996; González de la Rocha, 1999).¹⁷ The concept concerns payoffs and does not require altruism, local identity, conformity, or a larger stock of social capital.

Reciprocal exchange can compete with educational investment. Schooling takes children’s time away from local exchange and prepares them to leave the neighborhood, and these costs grow in adolescence, when children begin to contribute to their households (Lomnitz, 1977; González de

¹⁷In Ahmedabad, households offered subsidized housing on the city’s periphery lost access to informal insurance and borrowing, and many either did not move or later returned to central slums (Barnhardt et al., 2017).

la Rocha, 1999). In the Life Quality Survey, the association between trust in neighbors and years of schooling is more negative just east of the boundary than on the west side (Table 2, column 11; Table C.3). The sign is not obvious, because local networks can also support schooling (Coleman, 1988), and the experimental prediction does not require knowing which channel dominates.¹⁸

Experimental prediction. A minimal model delivers the prediction. Let ρ denote the return to investing in a child’s education, the same for all households, and C_i its private cost, which depends on educational provision P_ℓ , parental human capital h_i , and household resources y_i , the standard determinants examined in Section 3. Let $\Pi_i \geq 0$ denote the value of local exchange that household i gives up by investing, which is zero for households outside local exchange, and $a(s)$ the weight parents place on it when neighborhood salience is s . The household invests if

$$\rho - C_i \geq a(s) \Pi_i. \quad (2)$$

Making the neighborhood salient raises the weight from a_0 to a_1 and lowers the net payoff to education by $(a_1 - a_0) \Pi_i$, so educational demand falls where local exchange is valuable and not where it is absent. This is the first pre-registered hypothesis: neighborhood salience lowers demand for tutoring. The second, that the response is concentrated on the historically excluded side, follows if Π_i is larger there, as the evidence above suggests. Section 5 derives Π_i from the value of local exchange, the cost of participating, and the age of the child. The experiment varies salience, not Π_i itself.

4.3 Experimental Design

I test the prediction in a pre-registered incentivized choice experiment with parents and primary caregivers living within three kilometers of the historical boundary and with at least one child aged 6–20.¹⁹ The analysis includes 239 households, 193 east of the boundary and 46 west, an allocation pre-specified to increase power on the historically excluded side. Sampling was stratified by side: within each side, I selected census blocks with residents aged 6–20 in the 2020 census, and enumerators screened dwellings in those blocks. The mean distance to the boundary is 1.5 km and the median 1.7 km. The sample reproduces the adult schooling gap of its own

¹⁸The closest precedent is Munshi and Rosenzweig (2006), where caste-based occupational networks affect schooling choices by raising the value of remaining connected to traditional occupations. Here the network is defined by residence rather than caste, its benefits arise from reciprocal local exchange rather than occupational referrals, and the experiment varies neighborhood salience rather than network value.

¹⁹Fieldwork began on May 9, 2026 and was carried out by Descifra Investigación Estratégica using SurveyToGo. 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) on May 14, 2026. Twenty-three households were interviewed before that date, and excluding them leaves the results unchanged (Appendix D.1). Every eligible household contacted participated, and all received MXN 100 regardless of their choices; the 17 households interviewed before the payment was introduced on May 13 were compensated retroactively.

census blocks: respondents on the east side have 0.77 fewer years of schooling, against 0.75 among residents aged 15 and older of the same blocks in the 2020 census.

Neighborhood salience. The treatment changes the order of two tasks. In the mapping exercise, respondents view a satellite map centered on their residence and trace the area they consider their neighborhood (*colonia* or *barrio*), following McCartan et al. (2024). In the *Map-First* arm, they map immediately before the educational choice; in *Map-After*, immediately afterward. Assignment is randomized within side; the draw assigned 133 households to Map-First and 106 to Map-After. The mapping exercise mentions neither education nor the historical boundary, so the treatment changes whether respondents have just thought about their neighborhood, holding information and material incentives fixed (Figure D.1).

Educational demand. Respondents make eleven choices between one hour of online mathematics tutoring for one of their children and an electronic gift card from MXN 0 to MXN 250 in MXN 25 increments, following the multiple-price-list design of Jack et al. (2022). Row order and the side on which tutoring appears are randomized, one row is randomly selected to bind, and one in ten households receives the tutoring session or gift card implied by its choice.²⁰ The pre-registered outcomes are the share of choices in which tutoring is selected and latent willingness to pay for tutoring.

4.4 Effects of Neighborhood Salience on Educational Demand

Figure 5 shows that making the neighborhood salient reduces educational demand on the east side of the historical boundary. I estimate effects separately by side with enumerator and survey-day fixed effects, distance to the boundary with a separate slope on each side, controls for the randomized presentation of the choice list, and LASSO-selected controls, and I permute Map-First assignment within side for inference (Table D.4).²¹

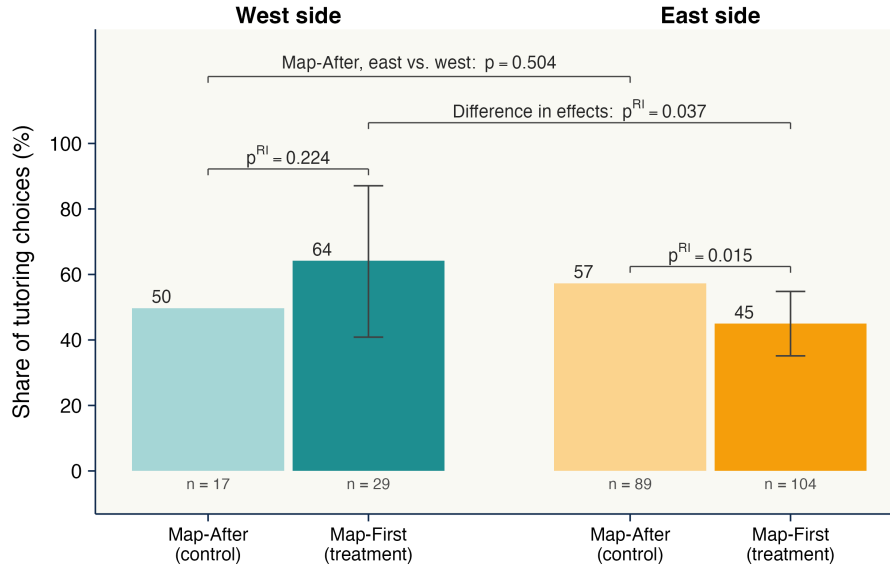
On the east side, Map-First lowers the tutoring share by 12.3 percentage points, from a Map-After mean of 0.573 to 0.450, about one fifth, and latent willingness to pay by MXN 75, from MXN 176 to MXN 102, about 42 percent. The reduction in willingness to pay equals about a quarter of the 2026 daily minimum wage of MXN 315. These effects correspond to 0.33 and 0.30 standard deviations, and restricting the sample to the historical Analco area produces nearly identical estimates.²² The difference with the west side has a randomization-inference p -value of 0.037 for

²⁰Before the task, respondents practice the same mechanism using a pen and quantities of marzipan candies and answer four comprehension questions; the mean score is 3.8 out of 4. Details are reported in Appendix D.

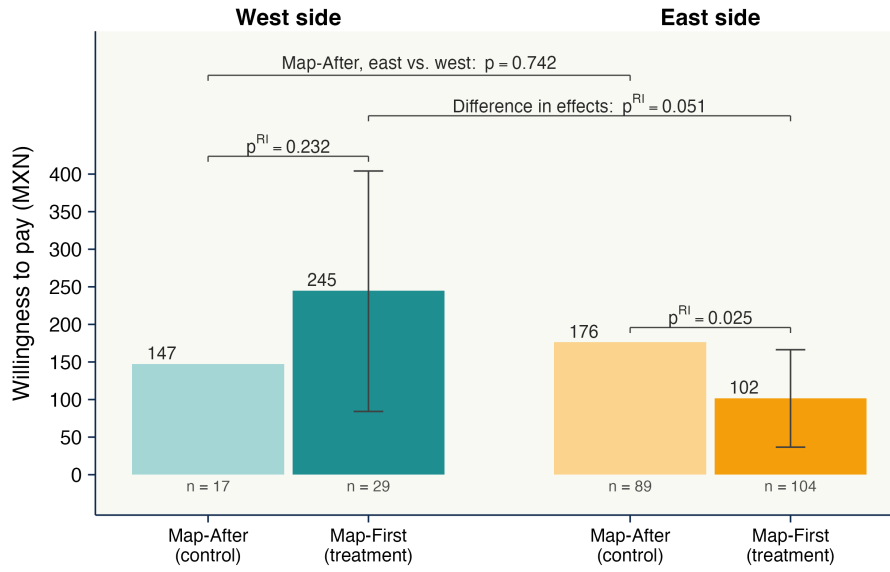
²¹Latent willingness to pay is recovered using the two-step estimator of Jack et al. (2022): a scaled random-effects probit of the eleven binary choices on gift-card value, followed by a regression of latent WTP on treatment assignment. Standard errors bootstrap both steps at the respondent level.

²²These magnitudes are close to the average effect of primes on overt behavior in the meta-analysis of Dai et al. (2023), $d = 0.38$ across 867 effects, and above the estimates that correct for publication bias, between 0.20 and 0.28.

Figure 5: Effects of neighborhood salience on educational demand



(a) Share of tutoring choices



(b) Willingness to pay

Notes: Light bars report unadjusted Map-After means; dark bars add the regression-adjusted Map-First effect to those means and are therefore not raw Map-First means. Panel (a) reports the share of the eleven choices in which tutoring is selected, and Panel (b) latent willingness to pay in MXN. Whiskers are 95 percent intervals obtained by inverting the randomization-inference test. Brackets report randomization-inference p -values for the side-specific effects and for their difference, and the p -value for the difference in unadjusted Map-After means across sides (Appendix D.2). Treatment labels are permuted 50,000 times within side, holding the number treated fixed. Regressions include enumerator and survey-day fixed effects, distance to the boundary, controls for choice-list presentation, and LASSO-selected controls. Labels below the bars report sample sizes by arm.

the tutoring share and 0.051 for willingness to pay. Without the prime, demand does not differ

Unlike most studies in that literature, the experiment was pre-registered and its outcome is an incentivized choice.

across sides: the Map-After means are similar on the east and west sides (Figure 5; $p = 0.50$).

On the west side, both estimates are positive and imprecise; with 46 households, and with Map-First households there having lived in their neighborhood 10.6 fewer years than Map-After households, I do not interpret their sign.

The experiment establishes that neighborhood salience changes educational demand on the east side. It does not identify why the response differs across sides, estimate the local cost Π_i , or explain the attendance gap in Section 3. Tutoring is a marginal educational investment valued against a fungible alternative: salience could lower the value of tutoring or raise the value of the gift card, for instance if the card helps meet obligations with neighbors. Both readings imply a higher opportunity cost of educational investment, and the experiment does not separate them.

4.4.1 Robustness and interpretation

The east-side result is not driven by imbalance, misunderstanding of the choice mechanism, a few respondents, or the measure of willingness to pay (Appendix D). The arms are balanced across more than thirty pre-treatment characteristics after false-discovery-rate adjustment, comprehension is nearly identical across arms, the estimate is stable when any single household is omitted, and a bounded area-under-demand measure gives the same result (Tables D.1 to D.3 and D.5 and Figure D.2).

The mapping task activates a compact area around the respondent's home rather than the historical divide. The median polygon covers about half a square kilometer, 88 percent contain the respondent's residence, only 11 percent cross the boundary, and descriptions refer mostly to physical extent and everyday use (Tables D.6 and D.7 and Figure D.3).

Experimenter demand does not match the spatial pattern of the effect. Thirty of 239 respondents inferred that the study linked neighborhoods to education, and Map-First raises this inference by 22 percentage points on the west side but not on the east side, where the behavioral effect appears. The east-side estimate is unchanged when these respondents are excluded or inference is controlled for, although the p -value of the difference between sides ranges from 0.04 to 0.11 across these post-treatment controls (Table D.8).

On the east side, parents who receive more help from neighbors demand less tutoring. Controlling for treatment assignment, one standard deviation more help received in the previous year is associated with a tutoring share 3.4 percentage points lower and a willingness to pay MXN 24 lower, statistically significant at ten percent (Figure D.4 in Appendix D.5).²³ The east-side effect is not concentrated among poorer or less-educated households; if anything, it is larger among households above the median income and with a parent who completed high school (Table D.9). Neither individual measures of local exchange nor their averages among other sampled households

²³Help received is measured after the experimental block, but the treatment does not change it ($p = 0.76$); this analysis was not pre-registered.

of the same *colonia* detectably moderate the response to salience. Because these measures barely vary across *colonias* within the east side, the sample cannot test whether the response varies with the density of local exchange.

A second randomized salience exercise targets the historical divide itself rather than the immediate neighborhood. Before respondents explain why a young person from their neighborhood might stop studying, a random half hear the expression *de la Calzada para allá*, which residents use to refer to the east side. On the east side, the expression lowers by 15 percentage points the share of explanations that refer to money or the need to work, and references to the neighborhood and to the family rise, although not significantly; on the west side, no explanation changes (Table D.10).²⁴ Because the treatment and the outcome differ from those of the tutoring experiment, this exercise is not a replication, but it shows that making the divide salient shifts how east-side residents explain schooling decisions toward the neighborhood.

4.5 The Local Social Environment Across the Boundary

Survey modules administered after the experimental block show that the clearest differences between the sides are in local exchange (Table 3). The comparisons control for Map-First assignment, respondent sex, age, schooling, and randomized question versions; they are observational and are not mediators of the treatment.

Local exchange and the cost of university. Analco respondents report receiving help from neighbors about twice as often during the previous year, and, conditional on work hours, time with neighbors accounts for a larger share of non-market time, about 7 percent against 4 percent on the west side.²⁵ Parents on the two sides report similar expected earnings and perceived university premia, but Analco respondents require an 18 percent higher salary at age 30, about MXN 5,300 per month, before they consider university worthwhile (Figure A.20). With similar perceived monetary returns, this gap is consistent with an additional nonpecuniary cost of university, although the question does not identify what that cost is. The premium is about 45 percent of the university premium these parents expect, close to the 42 percent drop in willingness to pay in the experiment, and under the assumption that the marginal parents' children would complete four more years of university, it implies a schooling gap of about 0.75 years, close to the 0.8 years in the census (Appendix D.8).

²⁴Not pre-registered. The explanations are classified with the same procedure as the descriptions of the neighborhood.

²⁵Some Analco respondents define the neighborhood as the area “where people know me and I know the neighbors” or “as far as people in the barrio know each other”; most answers on both sides refer to streets, landmarks, and routes. Appendix D.7 reports the answers and shares by side.

Table 3: Localized prosociality and alternative mechanisms

	Mean west side	Mean Analco	Analco coefficient	<i>q</i>
<i>Panel A. Localized prosociality (allocation exercises)</i>				
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
<i>Panel B. Localized network benefits and the cost of leaving</i>				
Times helped by neighbors, past 12 months	0.97	1.96	1.10** (0.44)	0.07
Share of non-market time with neighbors	0.041	0.073	0.040** (0.018)	0.07
Share of non-market time supporting children's studies	0.333	0.319	-0.015 (0.044)	0.78
Vignette: the stayer's family gets more help	0.52	0.60	0.17 (0.23)	0.63
Vignette: leaving for university strains relations	0.59	0.72	0.16 (0.25)	0.63
<i>Panel C. Aspirations and expectations</i>				
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
<i>Panel D. Returns to schooling and investment threshold</i>				
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
<i>Panel E. Perceived neighbor norms (pluralistic ignorance)</i>				
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
<i>Panel F. Stated attachment to the neighborhood</i>				
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-side 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 B, non-market time is the sum of weekly hours with neighbors, supporting the children's studies, and housework; the two share rows additionally control for weekly work hours and exclude respondents who report no non-market time. 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 sides, 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. Household characteristics of the sample are in Table D.11. All estimates are observational cross-boundary comparisons. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Giving to local counterparts. In an allocation exercise adapted from Enke et al. (2023) and Cappelen et al. (2025), respondents divide MXN 500 between a stranger from their own *barrio* and an equally well-off stranger elsewhere in Guadalajara. Analco respondents allocate about MXN 20 more to the neighborhood stranger, but overall giving does not differ, so the result reflects relative social weight rather than generosity.²⁶

What does not differ. Parents on both sides report similarly high aspirations for upper-secondary school and university, similar perceived returns to education, similar norms about working rather than studying, and similar feelings of belonging. Household income, size, parental schooling, relatives in the *colonia*, years of residence, and labor-market informality do not differ systematically in the experimental sample (Table D.11). These comparisons weigh against lower aspirations, different perceived returns, explicit norms, or socioeconomic composition as explanations for the experimental pattern, but they do not identify which feature of the environment generates it.

Reciprocal exchange and the culture of poverty. A familiar reading of lower educational demand in a poor, historically excluded neighborhood is the culture of poverty (Lewis, 1966), in which excluded groups share values such as fatalism and present orientation that lower investment in their children and pass to the next generation regardless of later opportunities. The evidence does not fit this reading. East-side parents report the same aspirations, perceived returns, and norms about working rather than studying as west-side parents. Their demand for tutoring falls only when the neighborhood has just been brought to mind, which a stable trait such as present orientation would not produce. The lower demand appears among households that receive more help from neighbors, not among the poorest, whereas Lewis (1966) characterized the culture of poverty by little organization beyond the family. The pattern is closer to what ethnographers of Mexican cities have described since the 1970s: households organized through exchange with neighbors, whose participation is valuable and costly to give up (Lomnitz, 1977; González de la Rocha, 1999).

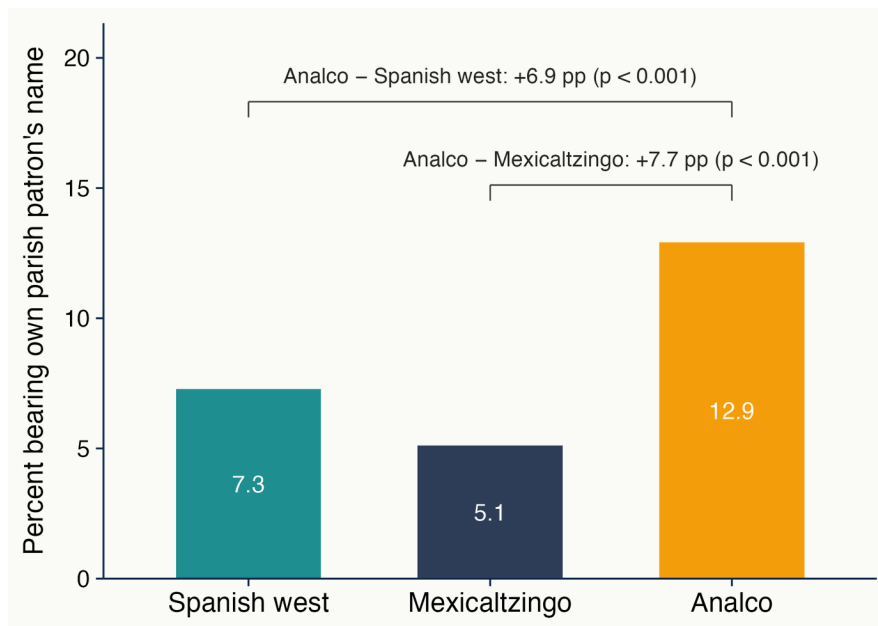
4.6 Historical Evidence of Locally Organized Social Life

Historical naming records indicate that social life on the eastern bank was locally organized well before the present. I use the civil *padrones* of 1821–22 and 1824 compiled by Anderson and Spike (2007), the 1930 census, and parish registers from 1850–1949; names are chosen within families but can reflect local religious and social environments (Fryer and Levitt, 2004; Abramitzky et al., 2020; Bazzi et al., 2020; Weigand et al., 2025; Kamel, 2026).

²⁶The equal split is modal on both sides; no west-side respondent gives the *barrio* stranger more than half, against 17 percent in Analco.

In 1821, 12.9 percent of men in Analco bore their parish patron’s name, against 7.3 percent in the Spanish west and 5.1 percent in Mexicaltzingo (Figure 6; Figure C.2). With name-token, cohort, and surname fixed effects, Analco exceeds the Spanish west by 6.9 percentage points and Mexicaltzingo by 7.7, while Mexicaltzingo and the Spanish west do not differ. Naming after the saint of the baptism day is equally common on both sides, and sons on the western side are more likely to bear their father’s name, so the difference is specific to the local parish, consistent with a parish-centered local identity (Figure C.3). Because Analco and Mexicaltzingo were both Indigenous settlements, the pattern is difficult to attribute to Indigenous status alone.

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



Notes: Bars report the share of men whose given name matches the patron saint of their own parish in the 1821 census (Anderson and Spike, 2007). Brackets report adjusted differences from a linear probability model at the token–individual level, weighted by the inverse 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. Figure C.2 maps the parishes, and Figure C.3 reports the baptism-day placebo.

The broader naming repertoire was also spatially distinctive: first names predict residence east of the creek for cohorts born from the late eighteenth through the nineteenth century, even within the ethnic categories of the 1821 census, and paternal surnames remain spatially distinctive in 1930 (Figure C.4). These records show historical differences in locally organized social life; they do not measure the value of local exchange or establish that a single trait persisted to the present.

4.7 Evidence Beyond Guadalajara

A similar descriptive relationship appears outside Guadalajara. Using Wave 7 of the World Values Survey, I construct a *neighborhood trust premium* in 35 countries: standardized trust in neighbors

minus average trust in family and personally known individuals. I relate it to years of schooling, educational mobility, and dropout before upper secondary, controlling for parental education and splitting countries by ethnic segregation (Alesina and Zhuravskaya, 2011) (Table E.5).

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*** (0.004)	-0.025*** (0.004)	-0.022*** (0.004)		-0.020*** (0.005)	-0.026*** (0.005)	-0.022*** (0.005)		0.014*** (0.004)	0.019*** (0.004)	0.014*** (0.004)	
Out-group trust		0.051*** (0.004)	0.047*** (0.005)			0.044*** (0.005)	0.039*** (0.006)			-0.041*** (0.005)	-0.034*** (0.005)	
Trust in close ties (family + known)			0.011** (0.005)				0.014*** (0.006)				-0.018*** (0.005)	
Neighborhood trust premium × Low ethnic segregation				-0.011** (0.005)				-0.014** (0.006)				0.005 (0.005)
Neighborhood trust premium × High ethnic segregation				-0.038*** (0.007)				-0.034*** (0.008)				0.027*** (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 how much trust is concentrated on neighbors relative to close ties 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.

The premium is negatively associated with schooling, particularly where ethnic groups are more segregated: a one-standard-deviation increase is associated with 0.038 standard deviations fewer years of schooling in high-segregation countries, against 0.011 in low-segregation ones (Table E.6). Trust in close ties and in distant groups are each positively associated with schooling; the negative association appears only when trust is concentrated on neighbors. These comparisons are descriptive: neighborhood trust is not reciprocal exchange, and segregation is not exogenous.²⁷

5 A Framework of Reciprocal Exchange

This section develops two simple models that combine elements of existing ones. A static model derives the local cost Π_i in (2) and rationalizes the experiment and the gap among children with similarly educated parents. A dynamic model rationalizes why the gap persists after educational provision converges. The value of reciprocal exchange comes from Kranton (1996), the network

²⁷A related exercise using the Facebook Social Connectedness Index finds lower schooling where a larger share of social ties is geographically local, but that measure captures network geography rather than reciprocal exchange (Figure E.1 and Table E.1).

that draws children away from education from [Munshi and Rosenzweig \(2006\)](#), and cultural transmission from [Bisin and Verdier \(2001\)](#); the ethnography of urban marginalization in Mexico gives each piece its content ([Lomnitz, 1977](#); [González de la Rocha, 1999](#)). The models organize the evidence; they are not estimated structurally. Appendix F states the results formally and proves them.

5.1 A Static Model of Educational Choice

Local exchange. Households obtain help, care, small loans, and information either through reciprocal exchange with neighbors or through outside channels such as markets, public services, and formal institutions. Let q_ℓ denote the share of households in neighborhood ℓ that participate in local exchange, and Z_ℓ the quality of outside channels. Following [Kranton \(1996\)](#), outside channels are thinner when more households exchange locally, so the surplus of local exchange over outside channels is

$$\Delta(q_\ell, Z_\ell) = R - Z_\ell (1 - q_\ell), \quad (3)$$

where R is the value of exchange with a neighbor. The surplus increases with participation and decreases with the quality of outside channels. Exclusion enters through Z_ℓ : it lowers the quality of outside channels materially, through public provision, and socially, through a group's standing in the institutions it must deal with. This is the substitution that [Lomnitz \(1977\)](#) documents, where exchange among neighbors provided the security that formal institutions did not.

Participation and education. Household i draws on local exchange with intensity $\beta_i \geq 0$, and $\theta_i \in \{0, 1\}$ indicates whether it participates. Participating costs $c(y_i)$, which decreases with household resources y_i : the household must remain available when neighbors need help and have resources to reciprocate ([González de la Rocha, 1999](#)). The household chooses whether to invest in the education of a child aged n , $e_i \in \{0, 1\}$. Education yields a return ρ that is the same for all households and has a private cost $C_i \equiv C(P_\ell, h_i, y_i) + \varepsilon_i$, where C decreases with educational provision P_ℓ , parental human capital h_i , and household resources, and ε_i is an idiosyncratic component. As in [Munshi and Rosenzweig \(2006\)](#), education draws the child away from the network: it takes the child's time and prepares the child to leave. Investing therefore reduces the value of participation by a share $\pi_n \in [0, 1)$ that increases with the child's age, because children begin to contribute to the household and its exchanges in adolescence ([González de la Rocha, 1994](#)). The household's payoff is

$$U_i = (\rho - C_i) e_i + \theta_i [(1 - \pi_n e_i) \beta_i \Delta(q_\ell, Z_\ell) - c(y_i)]. \quad (4)$$

The decision. The household participates if $\beta_i \Delta(q_\ell, Z_\ell) \geq c(y_i)$, so the poorest households may be unable to participate, as [González de la Rocha \(1999\)](#) documents. The cost of participating is paid whether or not the household invests, so it drops out of the investment decision. When

choosing, parents weigh the value of local exchange by $a(s) \in (0, 1]$, which increases with the salience s of the neighborhood, as in models of limited attention (Gabaix, 2019). Comparing $e_i = 1$ with $e_i = 0$ in (4) gives the condition in (2), $\rho - C_i \geq a(s) \Pi_i$, with

$$\Pi_i = \theta_i \pi_n \beta_i \Delta(q_\ell, Z_\ell). \quad (5)$$

The local cost of education is the part of the value of local exchange that education puts at risk. It is zero for households outside local exchange and larger where participation is dense, outside channels are poor, the household relies more on neighbors, and the child is older. Let $x_i \equiv \rho - C_i$ denote the net return, with distribution function F . Educational demand is $D_\ell(s) = \Pr(x_i \geq a(s)\Pi_i)$, and willingness to pay for a marginal investment, such as an hour of tutoring, is $W_i(s) = \omega [x_i - a(s)\Pi_i]$ for some $\omega > 0$.

Predictions. Let a_0 and $a_1 > a_0$ denote the weights in the Map-After and Map-First arms. The model delivers five predictions, stated formally in Appendix F.

First, making the neighborhood salient lowers demand among participating households and leaves it unchanged among the rest. The effect on average willingness to pay is $-\omega(a_1 - a_0) \mathbb{E}[\Pi_i]$, which is larger where more households participate and outside channels are worse. This is the experimental pattern if Π_i is larger on the east side, where residents receive more help from neighbors and trust formal institutions less.²⁸

Second, among participating households, demand decreases with the intensity β_i with which the household draws on local exchange, in both arms. East-side parents who receive more help from neighbors demand less tutoring (Section 4.4.1). The model also implies that the willingness-to-pay response to salience rises with β_i , an interaction the sample does not detect.

Third, households too poor to participate and households whose net return is far above the local cost respond little, so the response should not be concentrated among the poorest, as the experiment finds.

Fourth, with a common return and the same distribution of costs given parental human capital, attendance is lower where Π_i is larger, and the gap is concentrated among adolescents, for whom π_n is larger. Equalizing provision equalizes C_i but not Π_i , so it need not close the gap.

Fifth, parents with a larger local cost require a higher return, $C_i + a(s)\Pi_i$, before they consider education worthwhile, even when they perceive the same return. East-side parents require an 18 percent higher salary before they consider university worthwhile, with similar perceived returns.

5.2 Persistence Across Generations

The static model takes the participating share q_ℓ as given. To study persistence, I let the disposition to participate be transmitted from parents to children, following Bisin and Verdier (2001).

²⁸If the gift card can be used to meet obligations with neighbors, its value to participating households also rises with $a(s)$, which lowers the relative demand for tutoring in the same direction.

Transmission. A parent transmits its own disposition to the child with probability τ , chosen at a cost $\tau^2/2k$. Otherwise, the child adopts the disposition of a randomly met adult in the neighborhood, so transmission is local, as exchange is in [Lomnitz \(1977\)](#). A participating parent values a child who participates by $\lambda\Delta(q_\ell, Z_\ell)$, the value of keeping access to local exchange, and a non-participating parent values a child who relies on outside channels by $b(Z_\ell)$, which increases with their quality. Optimal efforts are $\tau_1 = k(1 - q_\ell)\lambda\Delta(q_\ell, Z_\ell)$ and $\tau_0 = kq_\ell b(Z_\ell)$: each type socializes more when its own type is scarce, and participating parents socialize more when local exchange is more valuable. The participating share evolves as

$$q_{\ell,t+1} - q_{\ell,t} = k q_{\ell,t} (1 - q_{\ell,t}) \Phi(q_{\ell,t}, Z_\ell), \quad \Phi(q, Z) \equiv (1 - q) \lambda\Delta(q, Z) - q b(Z). \quad (6)$$

Human capital is transmitted in parallel: a child's education becomes the parental human capital of the next generation and lowers its cost of education ([Cunha and Heckman, 2007](#)).

From participation to the gap. Participation determines the educational gap. For children aged n whose parents have human capital h , and abstracting from the participation cost and from differences in β_i , attendance in neighborhood ℓ is $1 - F_h(0) - q_{\ell t} [F_h(a_0\pi_n\beta\Delta_{\ell t}) - F_h(0)]$, where F_h is the distribution of net returns given h , a_0 is the weight on local exchange when the neighborhood is not made salient, and $\Delta_{\ell t} \equiv \Delta(q_{\ell t}, Z_{\ell t})$. The attendance gap between the west (W) and east (E) sides is therefore

$$G_t(h) = q_{Et} [F_h(a_0\pi_n\beta\Delta_{Et}) - F_h(0)] - q_{Wt} [F_h(a_0\pi_n\beta\Delta_{Wt}) - F_h(0)]. \quad (7)$$

The gap conditional on parental human capital rises with the participating share and with the value of local exchange on the east side, and it closes only if both converge.

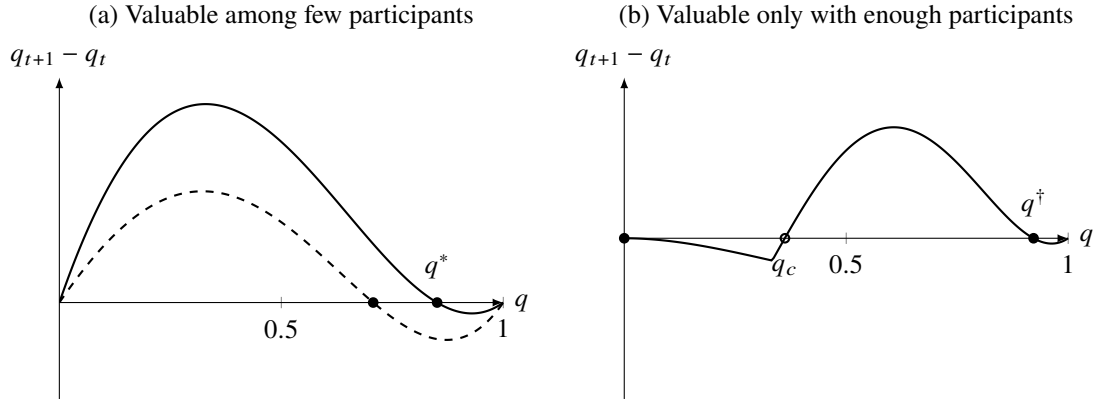
Two regimes. The law of motion (6) has two regimes, which depend on whether local exchange is valuable even when few neighbors participate, that is, on whether R exceeds Z_ℓ .

When outside channels are poor ($R > Z_\ell$), parents' incentives work as cultural substitution: the participating share converges to a unique stable level $q^*(Z_\ell)$ that decreases with the quality of outside channels (Figure 7, panel a). Exclusion raised q^* on the east side. When outside channels improve, the share declines toward a lower level, but only gradually, because each generation adjusts through socialization.

When outside channels are better ($R < Z_\ell$), local exchange is valuable only if enough neighbors participate, and the law of motion can have a tipping point q_c (Figure 7, panel b). A neighborhood above q_c converges to a stable state $q^\dagger$ with dense local exchange, while one below it loses participation. A neighborhood that entered this regime with dense participation, because exclusion had made it valuable, can remain there after outside channels converge.

Reading the gap over time. In 1930, the east side had worse educational provision and worse outside channels, so both the cost of education C_i and the local cost Π_i were higher there. As

Figure 7: Transmission of participation in local exchange



Notes: Change per generation in the participating share, q , from (6) with $\lambda = 1$ and $k = 1$. Panel (a) shows the regime in which local exchange is valuable among few participants ($R > Z$, with $R = 0.6$): the solid line uses poor outside channels ($Z = 0.2$, $b = 0.1$) and the dashed line better ones ($Z = 0.4$, $b = 0.2$); the stable level falls from 0.85 to 0.71. Panel (b) shows the regime in which local exchange is valuable only when enough neighbors participate ($R < Z$, with $R = 0.4$, $Z = 0.6$, $b = 0.03$): participation erodes below the tipping point $q_c \approx 0.36$ and converges to $q^\dagger \approx 0.92$ above it. Parameters are illustrative, not calibrated.

provision converged, it lowered the cost of education and improved outside channels, and the attendance gap fell from 10.6 to about 3 percentage points. What remains reflects the local cost, which declines only as participation and its value decline in (7). Parallel transmission of human capital also implies that parental human capital should absorb a larger share of the gap over time, as it does: from 11 percent in 1930, measured by the household head's literacy, to about a third in 2015, measured by parental schooling. The stability of the gap since 2000 is consistent with slow convergence in the first regime and with a stable state of dense participation in the second, which the data cannot distinguish.

Residential turnover. Transmission operates among residents who stay. If a fraction μ_ℓ of residents is replaced each period by entrants whose participating share is q_ℓ^M , then

$$q_{\ell,t+1} = (1 - \mu_\ell) \tilde{q}_{\ell,t+1} + \mu_\ell q_\ell^M, \quad (8)$$

where $\tilde{q}_{\ell,t+1}$ follows (6). Entrants who participate less lower q_ℓ and raise educational investment through (7). In the first regime, the neighborhood returns toward q^* after a shock; in the second, a large enough shock can push it below q_c . Section 6 studies such a shock, and a proxy for long-term residence is higher east of the boundary, although it does not measure turnover directly (Table B.13 and Appendix B.12).

What the model does not assume. The model does not assume that east-side households value education less: the return is common, and parents weigh it correctly. Demand is lower because investing puts an economic resource at risk, and what parents transmit is participation in an exchange whose value depends on outside channels and on the participation of neighbors.

In a culture of poverty (Lewis, 1966), by contrast, values that lower investment in children are transmitted regardless of later opportunities. In the model, persistence is conditional, which is why the gap narrowed as provision converged and why it should respond to changes in who lives in the neighborhood. This is a possible persistence mechanism, not an identified historical sequence: colonial exclusion may have changed provision, parental human capital, residential composition, and local social organization at the same time.²⁹

6 Residential Turnover and the 1992 Explosions

The framework in Section 5 implies that residential turnover can change both the human-capital composition of a neighborhood and the local environment in which educational choices are made. The 1992 explosions in Analco provide an unusual episode in which resident composition changed abruptly within a small part of the neighborhood.

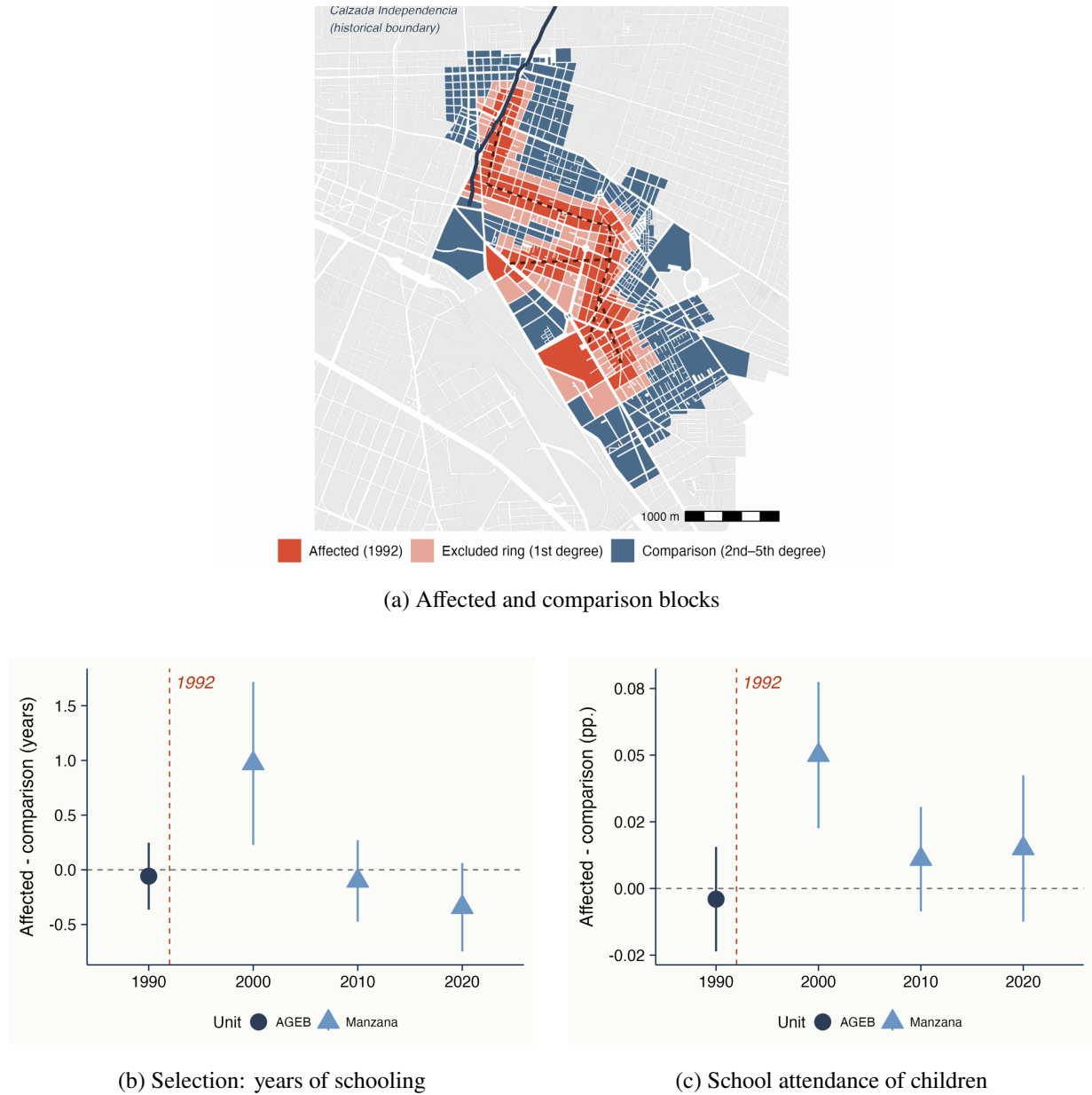
In April 1992, a series of explosions along the Reforma corridor destroyed roughly 1,500 buildings. The disaster displaced many former residents (Reguillo, 1998), and the rebuilt area was occupied largely by new households. I compare directly affected blocks with nearby unaffected blocks in Analco using the 2000, 2010, and 2020 censuses; a coarser 1990 comparison provides a pre-shock benchmark (Figure 8).

The first post-shock census shows positive rather than negative selection into the rebuilt area. In 2000, adults in affected blocks had almost one additional year of schooling relative to nearby comparison blocks. Children in those blocks were also about five percentage points more likely to attend school than children with similarly educated parents. Neither difference is visible in 2010 or 2020. At the coarser spatial scale available before the shock, the affected corridor also shows no pre-existing schooling advantage in 1990, although its adult population was already about one year older.

The census cannot determine why the initial advantage disappeared because it follows places rather than households. More educated entrants may have subsequently left, less educated households may have entered as the rebuilt housing aged, or both. The age composition of the blocks changed over the same decade, with adults in affected blocks 1.35 years older than in comparison blocks in 2000 and 0.86 years younger in 2010 ($p = 0.004$ for the change). Because the advantage of 2000 was in the completed schooling of adults, which cannot fall by a year within a decade among the same people, its disappearance suggests a new turnover of residents rather than adjustment among the same households. In the framework of Section 5, turnover changes the neighborhood

²⁹Local networks need not always compete with human capital; they may also support children's education (Coleman, 1988). Relocation away from established networks in Ahmedabad and Addis Ababa did not raise children's human capital (Barnhardt et al., 2017; Franklin, 2025), while children displaced by slum clearance in Santiago attained less schooling (Rojas-Ampuero and Carrera, 2026). The model isolates the margin studied here rather than claiming that local networks generally reduce educational investment.

Figure 8: The 1992 Reforma explosions: affected blocks, selection, and school attendance



Notes: Panel (a) maps blocks directly affected by the April 1992 explosions (red), defined as intersecting an 80-meter buffer around the collector's path (dashed line); immediate first-degree neighbors, excluded from the comparison (yellow); and second- through fifth-degree neighboring blocks used as the comparison group (teal). The dark line is Calzada Independencia. Panels (b) and (c) plot differences between affected and comparison areas by census wave. Panel (b) reports years of schooling among adults aged 18 and older; panel (c) reports school attendance among children aged 6–18 conditional on parental schooling. The 1990 estimate is a pre-shock comparison at the AGEB level; the 2000–2020 estimates use census blocks. Vertical bars show 95 percent confidence intervals. Appendix B.15 reports the construction of the affected area, comparison group, and specifications.

only when it is large enough relative to the forces sustaining local participation. A shock confined to a small set of blocks can therefore fade without changing the surrounding neighborhood.³⁰

The episode does not identify a causal place effect, and it does not reveal on which characteristics

³⁰This possibility is related to evidence that social networks can re-form after relocation. In Addis Ababa, networks among households moved to new housing sites recovered substantially within eight years (Franklin, 2025).

residents sorted. Its contribution is narrower. A sharp disruption of resident composition was followed initially by more educated residents and higher child attendance conditional on parental schooling, but these differences disappeared within the next decade as the composition of the blocks changed again. Residential sorting therefore appears as part of how the gap persists rather than as an alternative to it.

7 Conclusion

Nearly five centuries after Guadalajara's colonial boundary was drawn, it still predicts human capital gaps. Children on the historically excluded side are less likely to attend school, and adults there have fewer years of schooling. The gap has narrowed since 1930 but has not closed, even though schools on both sides are now comparable and the gap remains among children with similarly educated parents. In the experiment, asking parents to draw their neighborhood before an educational choice lowers their demand for tutoring on the excluded side. The prime is randomized among parents on the same side, so its effect does not reflect who lives there.

East-side parents report similar educational aspirations and expected returns to university, but they need a higher salary at age 30 to consider university worth it. They also receive more help from their neighbors, and among east-side parents, those who receive more help demand less tutoring. In 1821, naming patterns suggest that social life on the excluded side was already more closely tied to the local parish. These patterns are consistent with the ethnographic evidence of [Lomnitz \(1977\)](#) and [González de la Rocha \(1994\)](#). In the framework, following [Kranton \(1996\)](#), exchange with neighbors depends on continued participation, while schooling can draw children away from it. When parents pass this participation on to their children, lower investment in education can persist after school provision converges. This explanation does not require lower educational aspirations or a *culture of poverty* ([Lewis, 1966](#)).

We still know little about what households give up when they invest in education, or how much it is worth to them. It could be their children's time or the help they receive from neighbors. The experiment shows that thinking about the neighborhood changes the decision, but it does not tell us which consideration drives the response or how much of the attendance gap it explains. Nor are these households necessarily making a mistake: the help of their neighbors may be worth more to them than the tutoring. Future work should measure these costs and how they change as outside options improve.

Policies that ignore local ties may reach fewer of the households they target ([Bau et al., 2025](#); [Moscona et al., 2026](#)). Housing programs show how costly it can be to leave these relationships. Take-up ranged from 48 to 60 percent in relocation programs in the United States, India, and Brazil ([Chetty et al., 2016](#); [Barnhardt et al., 2017](#); [Belchior et al., 2026](#)). In India, many households that moved later returned after losing access to informal insurance and credit ([Barnhardt et al., 2017](#)). Schooling can strain the same relationships without any move, since a child in school helps the

neighbors less and is more likely to leave one day. Historical exclusion may persist through these relationships long after places converge in what they provide.

References

- Aaronson, Daniel, Daniel A. Hartley, and Bhashkar Mazumder**, “The Effects of the 1930s HOLC Redlining Maps,” *American Economic Journal: Economic Policy*, 2021, 13, 355–92. (cited on pp. 1 and 5).
- 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. 28).
- 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 8).
- , —, and —, “Institutions as a Fundamental Cause of Long-Run Growth,” 2005. (cited on p. 8).
- 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. 9).
- Akerlof, George A. and Rachel E. Kranton**, “Economics and identity,” *Quarterly Journal of Economics*, 2000, 115 (3), 715–753. (cited on p. 5).
- 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. 5).
- 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. 7).
- 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 17).
- 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. 30).
- Aliprantis, Dionissi**, “Making Our Neighborhoods, Making Our Selves: A Review Essay,” *Journal of Economic Literature*, 2023, 61 (3), 1172–1187. (cited on p. 6).
- Alsan, Marcella and Marianne Wanamaker**, “Tuskegee and the Health of Black Men,” *The Quarterly Journal of Economics*, 2018, 133 (1), 407–455. (cited on pp. 1 and 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. 9).
- 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. 15 and 21).
- 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. 28 and 29).
- 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).
- Austen-Smith, David and Roland G. Fryer**, “An Economic Analysis of “Acting White,”” *The Quarterly Journal of Economics*, 2005, 120 (2), 551–583. (cited on p. 5).
- Bagagli, Sara**, “The (Express)Way to Segregation: Evidence from Chicag,” Technical Report 2025. (cited on p. 9).
- 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 p. 137).
- 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 18).
- Banerjee, Abhijit V. and Andrew F. Newman**, “Information, the Dual Economy, and Development,” *The Review of Economic Studies*, 1998, 65 (4), 631–653. (cited on p. 5).
- Barnhardt, Sharon, Erica Field, and Rohini Pande**, “Moving to Opportunity or Isolation? Network Effects of a Randomized Housing Lottery in Urban India,” *American Economic Journal: Applied Economics*, 2017, 9 (1), 1–32. (cited on pp. 6, 21, 35, and 37).

- Bau, Natalie, Sara Lowes, and Eduardo Montero**, “Culture, Policy, and Economic Development,” 2025. Working paper, June 2025. (cited on p. 37).
- 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. 28).
- Becker, Gary S. and Nigel Tones**, “Human Capital and the Rise and Fall of Families,” *Journal of Labor Economics*, 1986, 4 (3, Part 2), S1–S39. (cited on pp. 2 and 17).
- 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. 19).
- Belchior, Carlos Alberto, Gustavo Gonzaga, and Gabriel Ulyssea**, “Who Benefits from Social Housing? Experimental Evidence from a Large-Scale Program in Brazil,” *American Economic Review: Insights*, 2026. Forthcoming. (cited on pp. 6 and 37).
- 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. 6).
- 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, 31, 32, 143, and 144).
- 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 8).
- Boltz, Marie, Karine Marazyan, and Paola Villar**, “Income Hiding and Informal Redistribution: A Lab-in-the-Field Experiment in Senegal,” *Journal of Development Economics*, 2019, 137, 78–92. (cited on p. 5).
- Borusyak, Kirill and Peter Hull**, “Nonrandom Exposure to Exogenous Shocks,” *Econometrica*, 11 2023, 91 (6), 2155–2185. (cited on p. 20).
- Bowles, Samuel and Herbert Gintis**, “Persistent parochialism: trust and exclusion in ethnic networks,” *Journal of Economic Behavior & Organization*, 9 2004, 55 (1), 1–23. (cited on p. 1).

- 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).
- Bryan, Gharad, Simon Franklin, Tigabu Getahun, and Sarah Winton**, “No Place Like Home? The Causal Effect of Forced Relocation from Central Addis Ababa,” 2026. Working paper, January 2026. (cited on p. 6).
- Buringh, E.**, “Urbanisation hub: The clio-infra database on urban settlement sizes: 1500–2000.,” 2018. (cited on p. 8).
- Bursztyn, Leonardo and Robert Jensen**, “How Does Peer Pressure Affect Educational Investments?,” *The Quarterly Journal of Economics*, 2015, 130 (3), 1329–1367. (cited on p. 5).
- **and —**, “Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure,” *Annual Review of Economics*, 8 2017, 9, 131–153. (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 and 14).
- 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. 8).
- Cappelen, Alexander W., Benjamin Enke, and Bertil Tungodden**, “Universalism: Global Evidence,” *American Economic Review*, 1 2025, 115, 43–76. (cited on p. 28).
- Carranza, Eliana, Aletheia Donald, Florian Grosset-Touba, and Supreet Kaur**, “The Social Tax: Redistributive Pressure and Labor Supply,” *Econometrica*, 11 2025, 93, 2273–2308. (cited on p. 5).
- Carrillo, Bladimir, Carlos Charris, and Wilman Iglesias**, “Moved to Poverty? A Legacy of the Apartheid Experiment in South Africa,” *American Economic Journal: Economic Policy*, 2023, 15 (4), 183–221. (cited on p. 6).
- 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. 11, 15, and 16).
- , —, and —, “Boundary Discontinuity Designs: Theory and Practice,” in R. Griffith, Y. Gorodnichenko, M. Kandori, and F. Molinari, eds., *Advances in Economics and Econometrics: Thirteenth World Congress*, Vol. 1, Cambridge University Press, 2026, chapter 2. To appear. (cited on pp. 2, 10, and 11).
- , **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, 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 pp. 6 and 37).
- , **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 p. 6).
- 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 p. 6).
- and **Lawrence F. Katz**, “Neighborhoods Matter: Assessing the Evidence for Place Effects,” *Journal of Economic Perspectives*, 11 2021, 35 (4), 197–222. (cited on p. 6).
- , **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 p. 9).
- Coleman, James S**, “Social Capital in the Creation of Human Capital,” *American Journal of Sociology*, 1988, 94 (1), S95–S120. (cited on pp. 1, 5, 22, and 35).
- Cunha, Flavio and James Heckman**, “The Technology of Skill Formation,” *American Economic Review*, 2007, 97 (2), 31–47. (cited on pp. 2, 17, and 33).
- Dai, Wenhao, Tianshu Yang, Benjamin X. White, Ryan Palmer, Emily K. Sanders, Jack A. McDonald, Melody Leung, and Dolores Albarracín**, “Priming Behavior: A Meta-Analysis of the Effects of Behavioral and Nonbehavioral Primes on Overt Behavioral Outcomes,” *Psychological Bulletin*, 2023, 149 (1–2), 67–98. (cited on p. 23).
- 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 8).

- de Arregui, Domingo Lázaro**, *Descripción de la Nueva Galicia* 1946. (cited on p. 8).
- Dell, Melissa**, “The Persistent Effects of Peru’s Mining Mita,” *Econometrica*, 2010, 78 (6), 1863–1903. (cited on pp. 1, 5, 8, and 11).
- Durlauf, Steven N.**, “Neighborhood Effects,” *Handbook of Regional and Urban Economics*, 1 2004, 4, 2173–2242. (cited on p. 6).
- Enke, Benjamin, 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. 28).
- Franklin, Simon**, “Enabled to Work: The Impact of Government Housing on Slum Dwellers in South Africa,” *Journal of Urban Economics*, 2020, 118, 103265. (cited on p. 6).
- , “The Demand for and Impacts of Government Housing: Evidence from Ethiopian Lotteries,” 2025. Working paper, Queen Mary University of London, May 2025. (cited on pp. 6, 35, and 36).
- Fryer, Roland G. and Steven D. Levitt**, “The Causes and Consequences of Distinctively Black Names,” *The Quarterly Journal of Economics*, 2004, 119 (3), 767–805. (cited on p. 28).
- Gabaix, Xavier**, “Behavioral Inattention,” in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 2, North-Holland, 2019, pp. 261–343. (cited on p. 32).
- Galor, Oded**, “Inequality, Human Capital Formation, and the Process of Development,” in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 4, Elsevier, 2011, chapter 5, pp. 441–493. (cited on pp. 2 and 17).
- **and Joseph Zeira**, “Income Distribution and Macroeconomics,” *The Review of Economic Studies*, 1993, 60 (1), 35–52. (cited on pp. 2 and 17).
- Genicot, Garance and Debraj Ray**, “Aspirations and Economic Behavior,” *Annual Review of Economics*, 8 2020, 12, 715–746. (cited on p. 17).
- Giuliano, Paola and Nathan Nunn**, “Understanding Cultural Persistence and Change,” *Review of Economic Studies*, 2021, 88 (4), 1541–1581. (cited on p. 5).
- Goldberg, Jessica**, “The Effect of Social Pressure on Expenditures in Malawi,” *Journal of Economic Behavior & Organization*, 2017, 143, 173–185. (cited on p. 5).
- 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. 8).

- González de la Rocha, Mercedes**, *The Resources of Poverty: Women and Survival in a Mexican City* Studies in Urban and Social Change, Oxford: Blackwell, 1994. (cited on pp. 1, 3, 20, 31, and 37).
- , “La reciprocidad amenazada: un costo más de la pobreza urbana,” in Rocío Enríquez Rosas, ed., *Hogar, pobreza y bienestar en México*, Guadalajara: ITESO, 1999, pp. 13–36. (cited on pp. 1, 20, 21, 28, and 31).
- Goraya, Sampreet Singh, Suanna Oh, and Yogita Shamdasani**, “Absent but Not Idle: The Social Roots of Demand for Flexible Work,” September 2025. Working paper. (cited on p. 5).
- 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).
- 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*, 1^o ed., Mexico: UNAM, Instituto de Investigaciones Históricas, 2010, pp. 105–122. (cited on p. 7).
- Hanlon, W.Walker and Stephan Heblich**, “History and urban economics,” *Regional Science and Urban Economics*, 5 2022, 94, 103751. (cited on p. 5).
- 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. 1, 5, and 9).
- Hornbeck, Richard and Suresh Naidu**, “When the Levee Breaks: Black Migration and Economic Development in the American South,” *American Economic Review*, 2014, 104 (3), 963–990. (cited on p. 6).
- , **Guy Michaels, and Ferdinand Rauch**, “Identifying Agglomeration Shadows: Long-run Evidence from Ancient Ports,” Technical Report 2025. (cited on p. 9).
- Huillery, Elise**, “History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa,” *American Economic Journal: Applied Economics*, 4 2009, 1, 176–215. (cited on pp. 1 and 5).
- 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 p. 23).
- Jakiela, Pamela and Owen Ozier**, “Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies,” *Review of Economic Studies*, 2016, 83 (1), 231–268. (cited on p. 5).
- 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 estrategias,” Vol. 6, COMECOSO and CUCSH-UdeG, 2016, pp. 77–106. (cited on pp. 2 and 10).
- Jaramillo-Molina, Máximo Ernesto**, ““De la Calzada para allá”: desigualdad, segregación y estigmatización en el Área Metropolitana de Guadalajara, México,” *Laboratorio*, 2022, 32 (2), 11–55. (cited on pp. 2, 10, and 20).
- 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. 28).
- Katz, Lawrence F.**, “Presidential Address: Neighborhood Effects and Missing Markets for Opportunity,” *American Economic Review*, 2026, 116 (9), 3197–3222. (cited on p. 6).
- Keele, Luke and Rocío Titiunik**, “Geographic Boundaries as Regression Discontinuities,” *Political Analysis*, 1 2015, 23 (1), 127–155. (cited on p. 11).
- , **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. 11).
- 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. 6).
- Knight, Alan**, *Mexico. The colonial era*, Cambridge University Press, 2002. (cited on p. 7).
- Kranton, Rachel E.**, “Reciprocal Exchange: A Self-Sustaining System,” *American Economic Review*, 1996, 86 (4), 830–851. (cited on pp. 1, 4, 5, 21, 30, 31, 37, and 140).
- Lee, David S. and Thomas Lemieux**, “Regression Discontinuity Designs in Economics,” *Journal of Economic Literature*, 2010, 48 (2), 281–355. (cited on p. 11).

- 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. 8).
- Lewis, Oscar**, “The Culture of Poverty,” *Scientific American*, 1966, 215 (4), 19–25. (cited on pp. 4, 28, 35, and 37).
- Lin, Jeffrey and Ferdinand Rauch**, “What future for history dependence in spatial economics?,” *Regional Science and Urban Economics*, 5 2022, 94, 103628. (cited on p. 5).
- Lomnitz, Larissa Adler**, *Networks and Marginality: Life in a Mexican Shantytown*, New York: Academic Press, 1977. Translated by Cinna Lomnitz. (cited on pp. 1, 3, 20, 21, 28, 31, 33, and 37).
- 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).
- **and Eduardo Montero**, “The Legacy of Colonial Medicine in Central Africa,” *American Economic Review*, 4 2021, 111 (4), 1284–1314. (cited on pp. 1 and 5).
- **, 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. 8).
- 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. 6).
- Maloney, William F. and Felipe Valencia-Caicedo**, “The Persistence of (Subnational) Fortune,” *The Economic Journal*, 12 2016, 126 (598), 2363–2401. (cited on p. 8).
- Matías, Alejandro Solís**, *Analco*, Gobierno del Estado de Jalisco, 1986. (cited on p. 7).
- **, “San José de Analco en 1907,”** *Cuadernos de Estudios Jaliscienses*, 1992. (cited on pp. 7 and 8).
- 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. 3, 6, and 23).
- 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).

- Moscona, Jacob, Nathan Nunn, and James A. Robinson**, “Searching for Fish in Trees? Economic Development when Context Matters,” NBER Working Paper 34810, National Bureau of Economic Research 2026. (cited on p. 37).
- Müller, Ulrich K. and Mark W. Watson**, “Spatial Unit Roots and Spurious Regression,” *Econometrica*, 9 2024, 92 (5), 1661–1695. (cited on p. 19).
- Munshi, Kaivan**, “Community Networks and the Process of Development,” *Journal of Economic Perspectives*, 2014, 28 (4), 49–76. (cited on pp. 1 and 5).
- **and Mark Rosenzweig**, “Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy,” *American Economic Review*, 2006, 96 (4), 1225–1252. (cited on pp. 1, 4, 5, 22, and 31).
- **and —**, “Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap,” *American Economic Review*, 2016, 106 (1), 46–98. (cited on pp. 1 and 5).
- 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 p. 6).
- 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. 5).
- Portes, Alejandro**, “Social Capital: Its Origins and Applications in Modern Sociology,” *Annual Review of Sociology*, 8 1998, 24, 1–24. (cited on pp. 1 and 5).
- 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. 1, 5, 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. 3, 10, 20, and 35).

- Rojas-Ampuero, Fernanda and Felipe Carrera**, “Sent Away: Displacement, Neighborhoods, and Children’s Outcomes under Slum Clearance Policies,” 6 2026. Working paper. (cited on pp. 6 and 35).
- Sacerdote, Bruce**, “Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?,” in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 3, Elsevier, 2011, pp. 249–277. (cited on p. 5).
- 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 p. 6).
- Sen, Amartya**, “Social Exclusion: Concept, Application, and Scrutiny,” Social Development Papers 1, Asian Development Bank, Manila, Philippines 6 2000. (cited on pp. 1 and 9).
- Shayo, Moses**, “Social Identity and Economic Policy,” *Annual Review of Economics*, 8 2020, 12 (1), 355–389. (cited on p. 5).
- Squires, Munir**, “Kinship Taxation as an Impediment to Growth: Experimental Evidence from Kenyan Microenterprises,” *Economic Journal*, 2024, 134, 2558–2579. (cited on p. 5).
- 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. 9).
- , *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. 6).
- 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 pp. 1 and 5).
- Valenzuela-Casasempere, Pablo**, “Displacement and Infrastructure Provision: Evidence from the Interstate Highway System,” Technical Report 2025. (cited on p. 9).
- Voth, Hans-Joachim**, “Persistence – myth and mystery,” *The Handbook of Historical Economics*, 1 2021, pp. 243–267. (cited on p. 5).
- Wantchekon, Leonard, Marko Klačnja, and Natalija Novta**, “Education and Human Capital Externalities: Evidence from Colonial Benin,” *The Quarterly Journal of Economics*, 5 2015, 130, 703–757. (cited on pp. 1 and 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. 28).

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. 64).

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. Appendix A.4 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. Appendix A.5 reports the 1882 cadaster and the early capitalization of exclusion into land values, and Appendix A.6 traces the persistence of the urban form from the colonial creek to present-day Calzada Independencia. Appendix A.7 reports the 1821 socio-demographic comparison across banks and Appendix A.8 the skin-tone evidence. Appendix A.9 reports the realized university wage premium.

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.

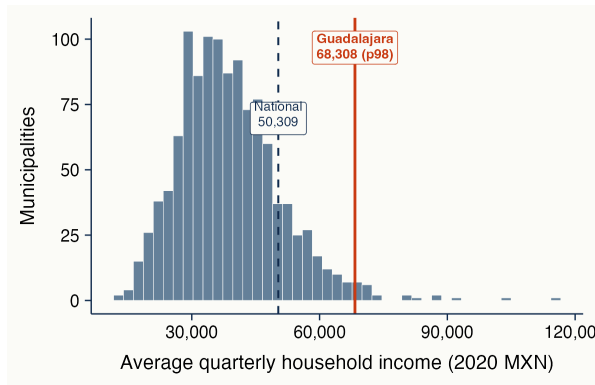
A.2 School attendance over time and where it varies

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.

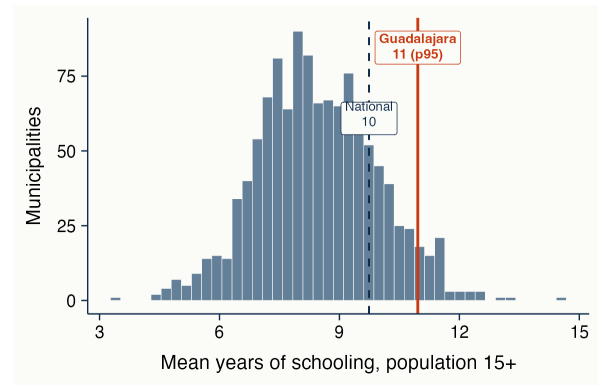
Figure A.3 asks where that variation lives. Writing y for whether a young person aged 12–17 attends school, the variance splits exactly into three nested parts: differences between municipalities, differences between households of the same municipality, and differences between siblings of the same household. 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

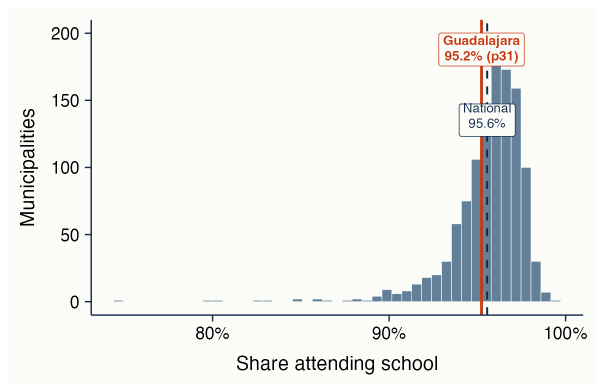
Figure A.1: Guadalajara in the national distribution of income, schooling, and attendance (2020)



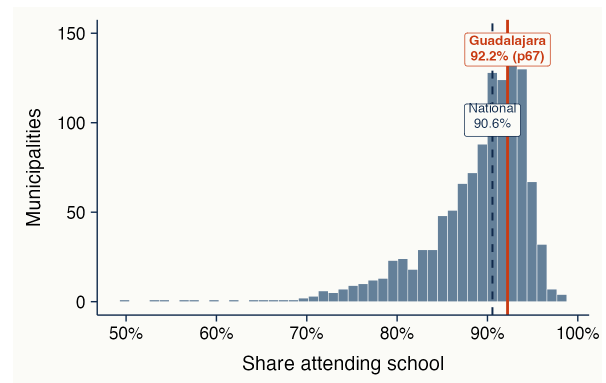
(a) Quarterly household income



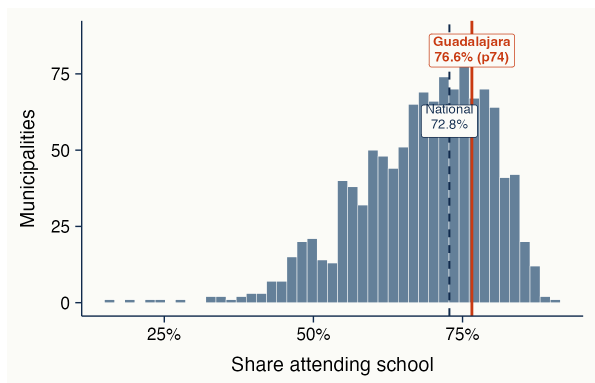
(b) Years of schooling, ages 15+



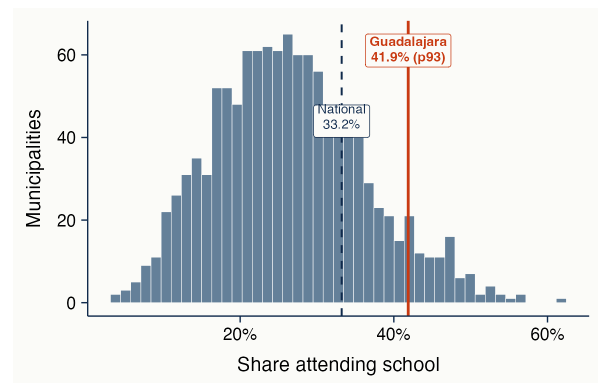
(c) School attendance, ages 6–11



(d) School attendance, ages 12–14



(e) School attendance, ages 15–17

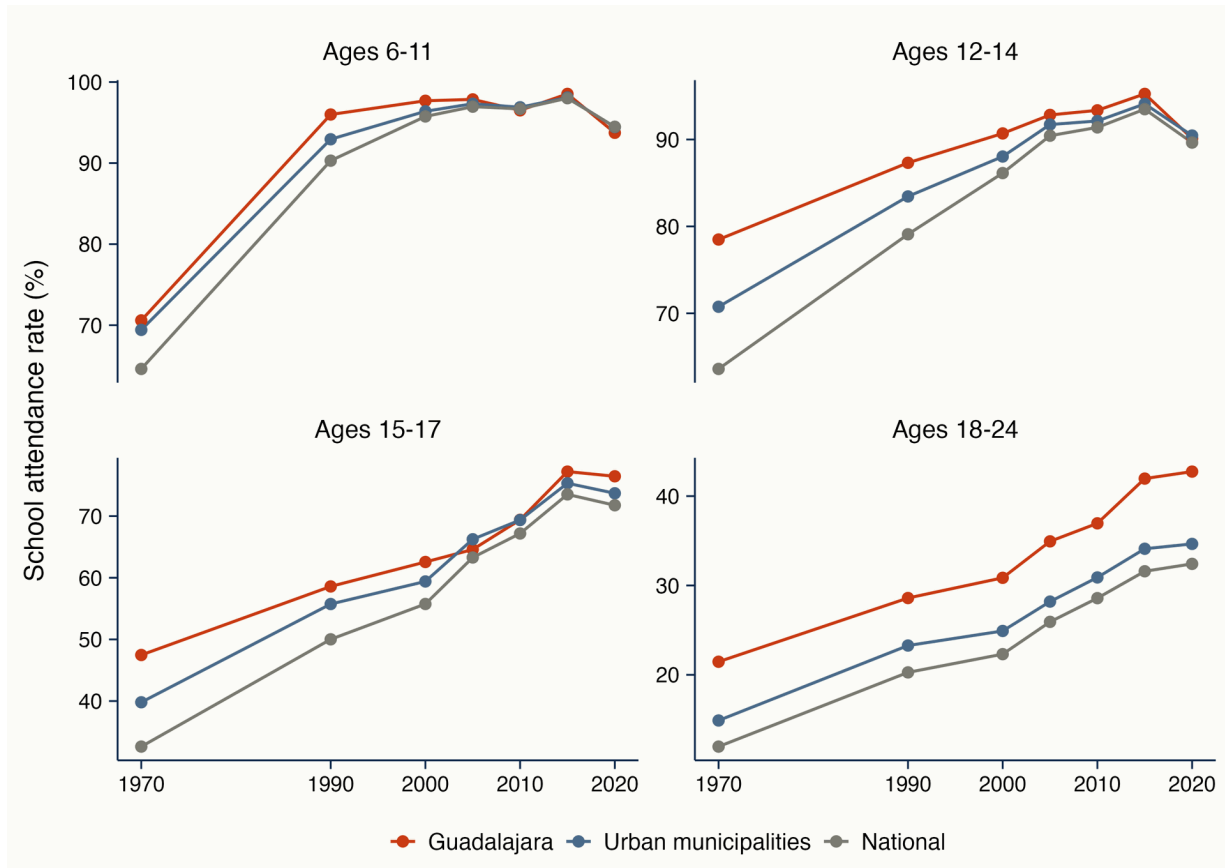


(f) School attendance, ages 18–24

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.

household component. Age, sex, the schooling of the household head or spouse, birth order 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

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



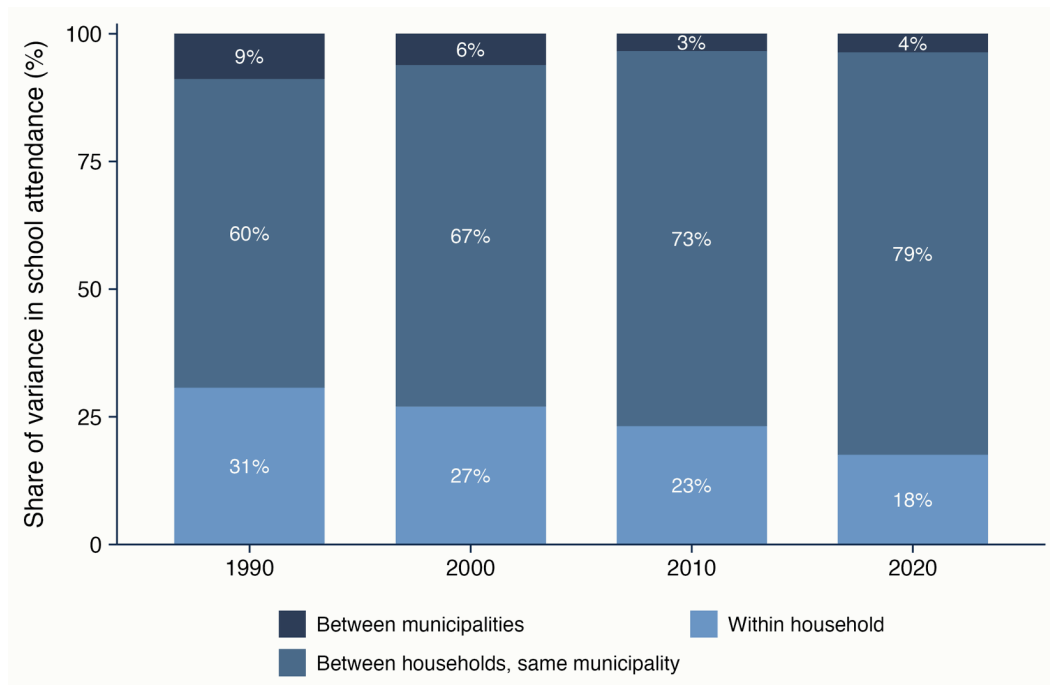
Notes: Share of individuals in each age group reported as attending school, computed from the IPUMS International samples of the Mexican censuses of 1970, 1990, 2000, 2010 and 2020 and the intercensal counts of 2005 and 2015, weighted by person weights. “Urban municipalities” are the harmonized municipalities 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.

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.

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



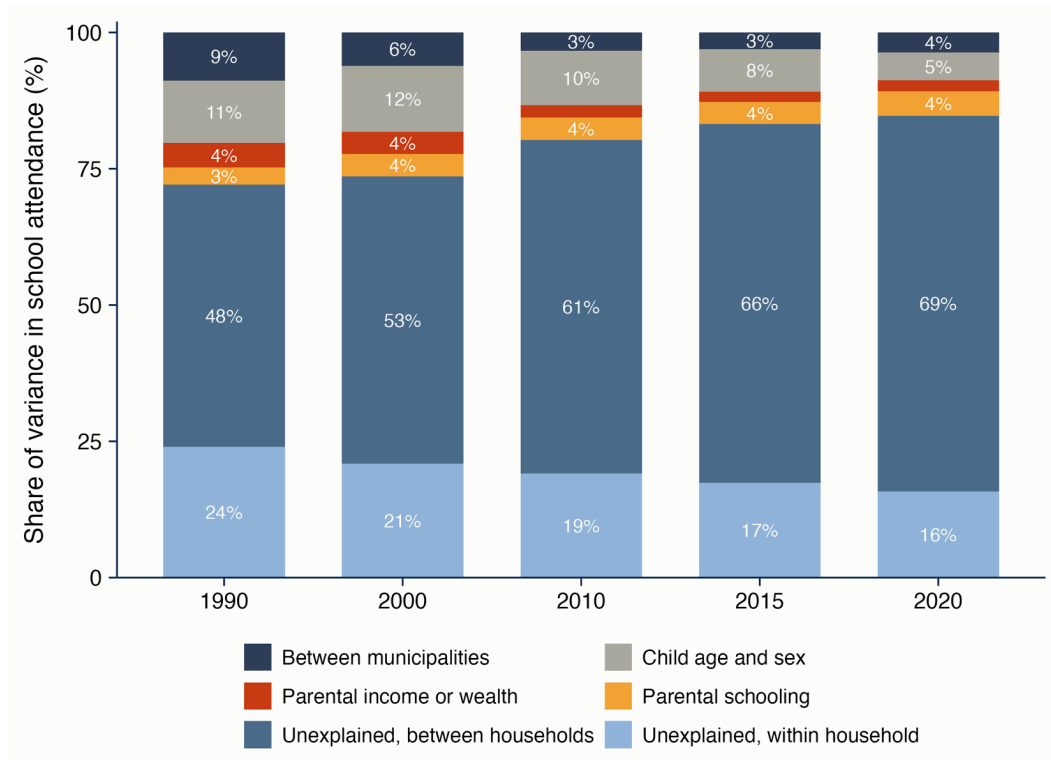
Notes: Nested decomposition of the variance of an indicator for attending school among individuals aged 12–17 who are children of the household head, using the IPUMS International samples of the 1990, 2000, 2010 and 2020 Mexican censuses and weighting by person weights. The three components are $\text{Var}_m(\bar{y}_m)$, $\mathbb{E}_m[\text{Var}_h(\bar{y}_h)]$ and $\mathbb{E}_h[\text{Var}(y)]$, 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.

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.8 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 Mexico state and home to Tenochtitlán — concentrated densely settled urban polities. The Atemajac Valley (Panel c), where Guadalajara would later be

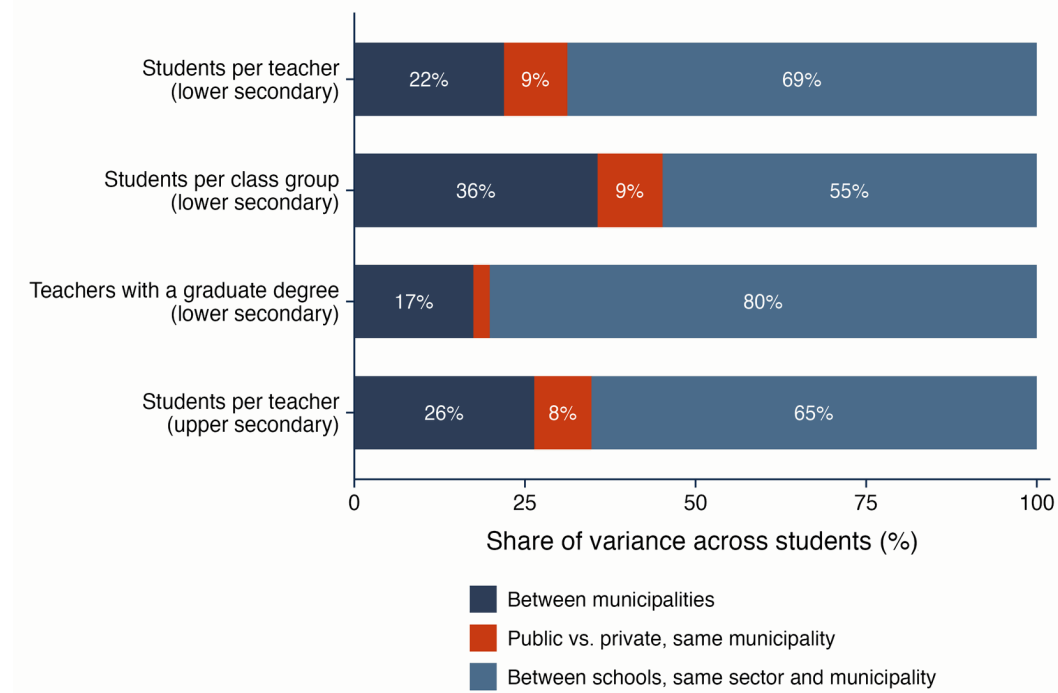
Figure A.4: Parental resources and the variance in school attendance, 1990–2020



Notes: 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 censuses and the 2015 Intercensal Survey, weighted by person weights. The between-municipality component is as in Figure A.3. The within-municipality variance is split with municipality fixed effects into the contributions of four blocks of indicators, allocated by Shapley values over all orderings: the child’s age by sex; parental income or wealth (deciles of the head’s and spouse’s labor earnings per household member, and deciles of the first principal component of electricity, piped water, sewage, gas or electric cooking fuel, television, radio, home ownership, and rooms per person); and parental schooling (years of the more schooled of head and spouse). Labor earnings of other members are excluded because they include the earnings of working children. The unexplained remainder is split into its between-household and within-household parts. Components sum to the total variance.

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.9 reports population estimates from Buringh (2018) for Mexico City and Guadalajara from the 16th to 18th centuries, showing that Guadalajara grew slowly 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.

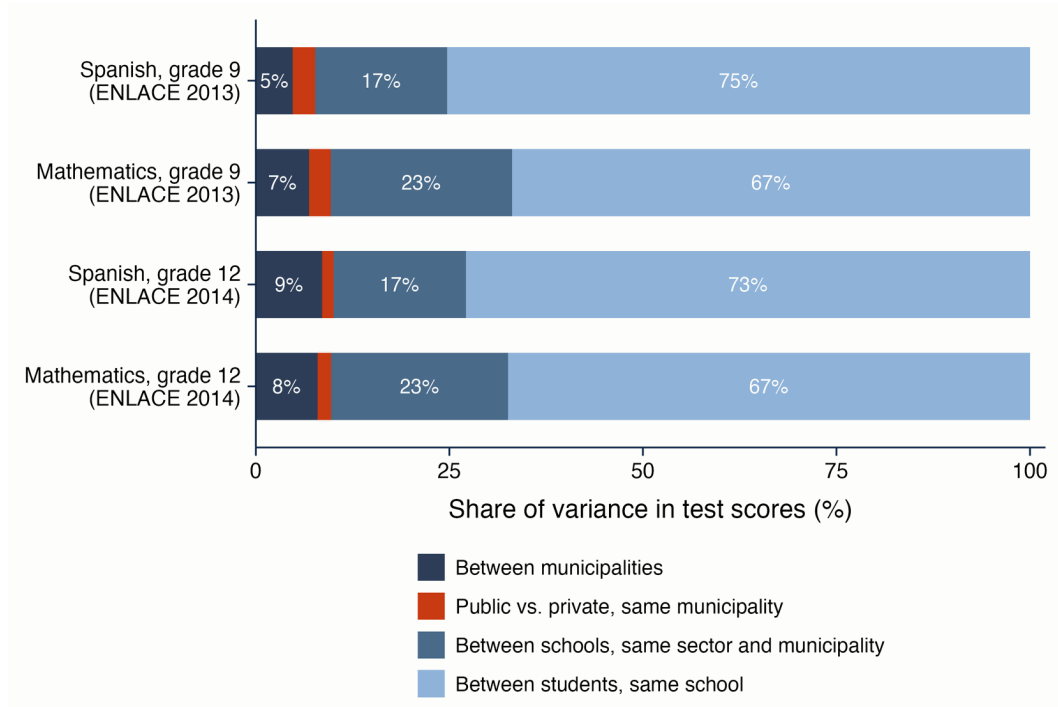
Figure A.5: Where the variance in school supply lies



Notes: Nested decomposition of the national variance of school characteristics across students into differences between municipalities, between public and private schools of the same municipality, and between schools of the same sector and municipality. Each school is weighted by its enrollment. Private schools hold 9.4 percent of lower-secondary enrollment. Lower secondary: *Formato 911*, 2018–2019, 34,195 schools; upper secondary: *Formato 911* for general and technological high schools, 2018–2019. Student–teacher and student–group ratios above the 99.5th percentile are set to missing as likely capture errors. Attending a private school is a household choice, and differences between schools combine differences in supply with the sorting of students across schools.

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.10 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 Boundary Discontinuity (BD) 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

Figure A.6: Where the variance in test scores lies



Notes: Nested decomposition of the national variance of individual ENLACE scores into differences between municipalities, between public and private schools of the same municipality, between schools of the same sector and municipality, and between students of the same school. Private schools are the *Particular* modality in grade 9 and the private *sostenimiento* in grade 12. Grade 9: ENLACE 2013, 1,589,219 students, excluding those flagged for copying. Grade 12: ENLACE 2014 for upper secondary, 972,272 students, excluding those who answered fewer than half the items in either subject, as in official averages; this file carries no copying flag.

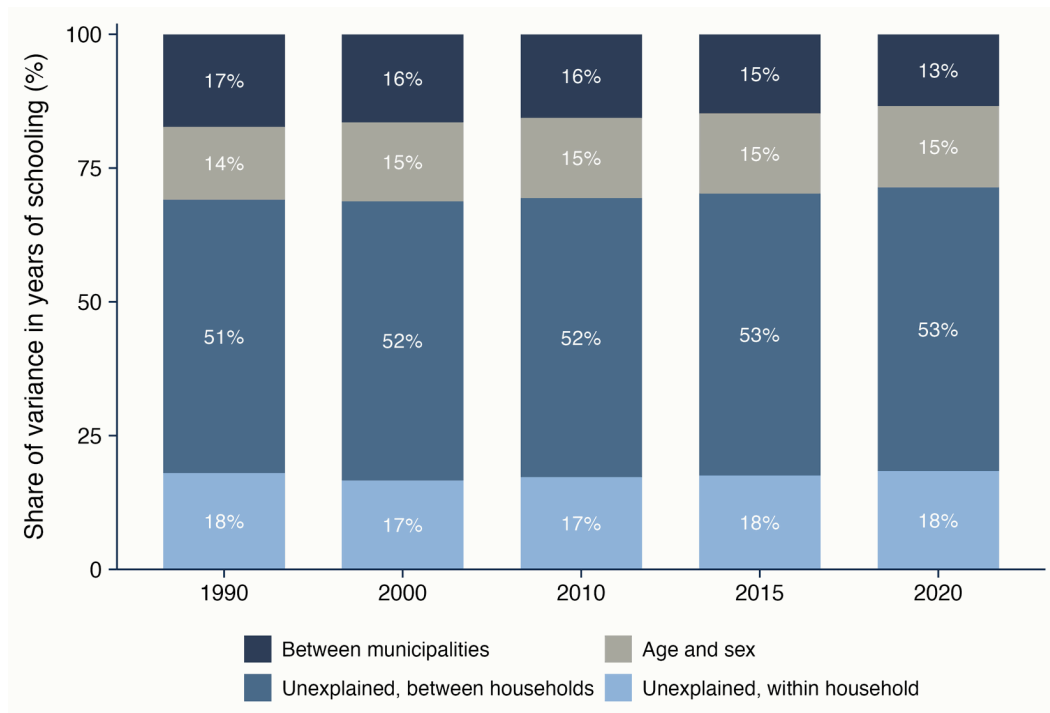
normals for 1902–2011, so smoothness across a three-kilometer window is in part a property of the interpolation rather than an independent finding, and two of them leave intervals reaching about 0.45 standard deviations. No layer measures hydrology, and the one hydrological feature at the boundary is 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 Named institutions on each bank, 1800–1908

This subsection documents the evidence behind Section 2.2 and Figure A.11: 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 (Figure A.12). The 1800 sheet is a facsimile of a plan of the city “as it

Figure A.7: Where the variance in adult years of schooling lies, 1990–2020



Notes: Decomposition of the variance of years of schooling among individuals aged 25 and older, using the IPUMS International samples of the 1990, 2000, 2010 and 2020 censuses and the 2015 Intercensal Survey, weighted by person weights. The between-municipality component uses harmonized municipalities. The within-municipality variance is split with municipality fixed effects into the contribution of age-by-sex indicators and an unexplained remainder, which is divided into its between-household and within-household parts. Components sum to the total variance. Parental schooling and resources are not included because the parents of adults rarely live in the same household; household income and wealth are not included because for adults they depend on their own schooling.

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

Table A.1: Balance in locational fundamentals

	BD estimate	SE	95% CI	<i>p</i> -value	FDR <i>q</i> -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 BD 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.

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 BD design.

Table A.2 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 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.2 reports the located entries by category, wave, and bank.

Table A.2: Located institutions by category, wave, and bank

Category	1800		1863		1890		1908	
	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.

A.5 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

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 (*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.6 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.15, panel (a) reproduces a 1908 map: nearly four centuries after the foundation, the city's footprint remained largely contained within the colonial core, 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.

From creek to avenue. Figure A.16 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.

The Calzada as a present-day connector, not a barrier. Figure A.17 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 BD design exploits a historical line whose physical form has changed repeatedly, rather than a contemporary infrastructural divide imposing its own treatment effect.

A.7 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.18 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 BD design 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.13). 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 than one series.

A.8 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.3: 1821 socio-demographic gaps across the boundary, with and without a control for *calidad*

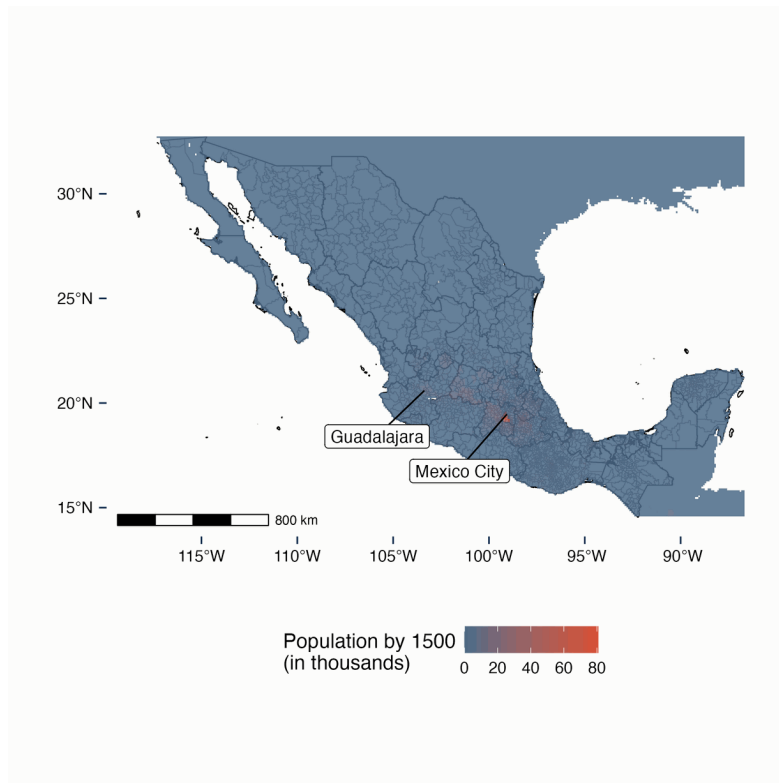
	(1) Full sample	(2) <i>Calidad</i> subsample	(3) <i>Calidad</i> controlled	Observations
Don/Doña	-0.374*** (0.016)	-0.337*** (0.021)	-0.176*** (0.022)	12,872 / 6,650
Interracial marriages	0.113*** (0.018)	0.113*** (0.018)	0.103*** (0.018)	2,869 / 2,869
Journeyman	0.211*** (0.019)	0.195*** (0.027)	0.130*** (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.18 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.18. 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.

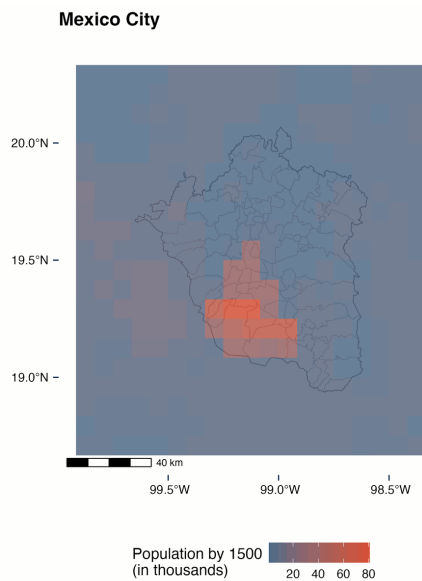
A.9 Realized university wage premium, 1990–2020

Figure A.20 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 4 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.

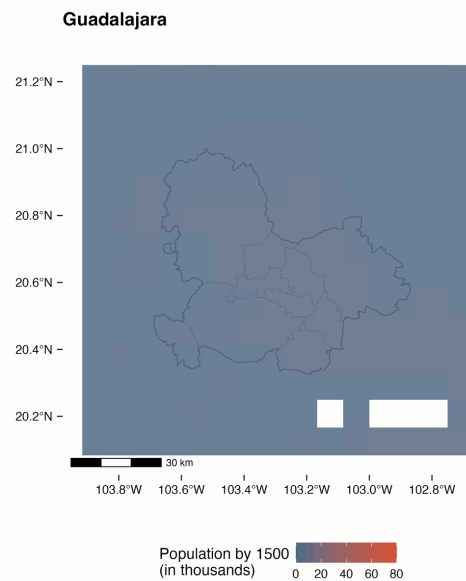
Figure A.8: Pre-colonial population density estimates in Mesoamerica



(a) Pre-colonial population density across Mexico



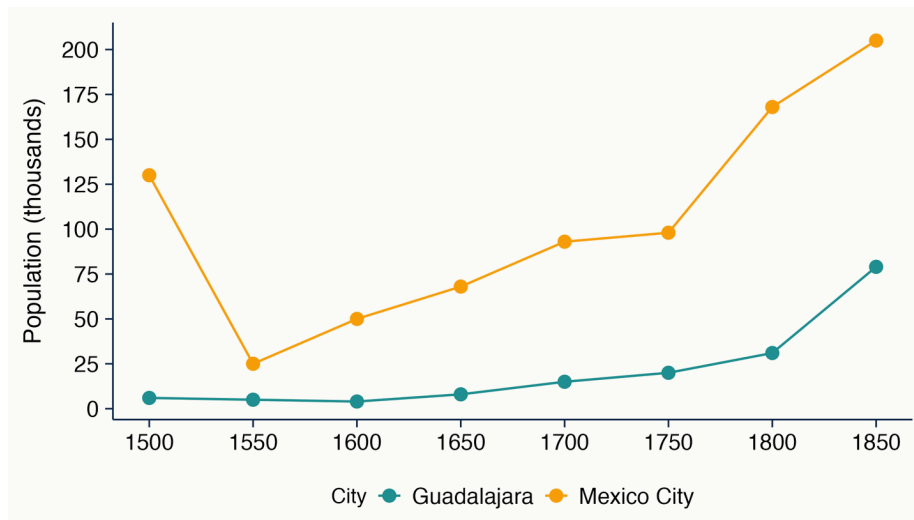
(b) Anahuac Valley (Central Mexico)



(c) Atemajac Valley (Guadalajara region)

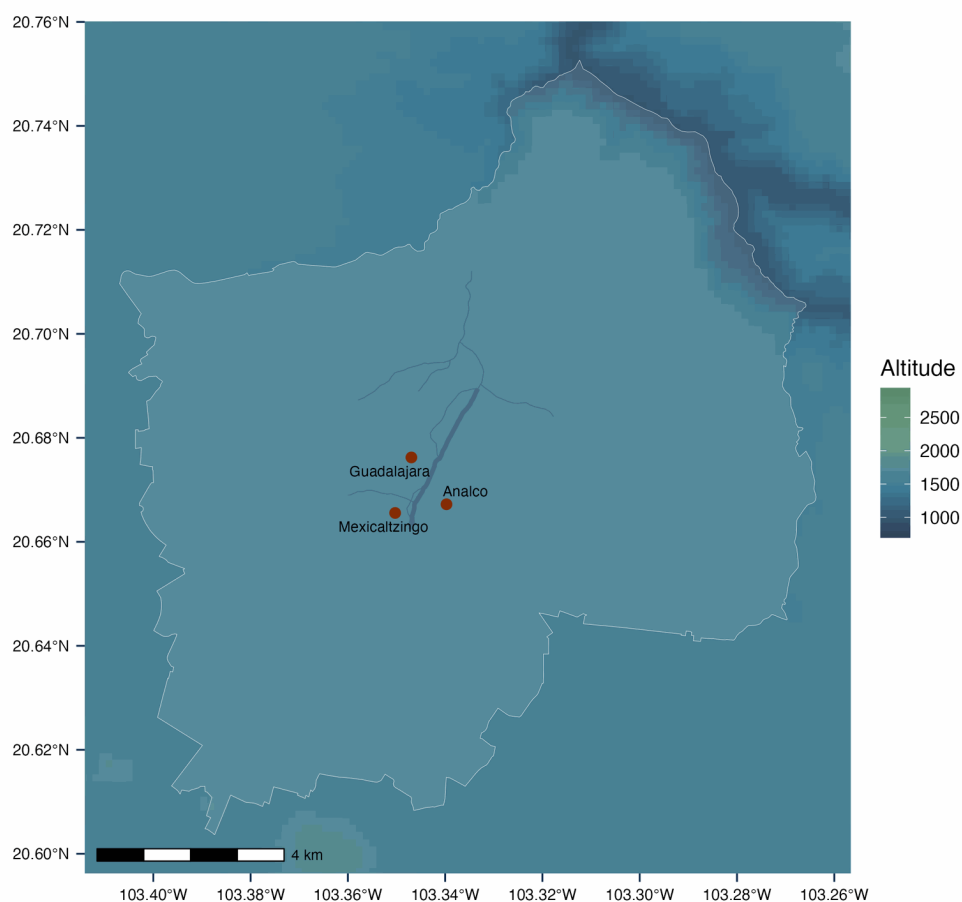
Notes: Panel (a) shows pre-colonial population density across present-day Mexico circa 1500 CE, based on the historical land-use reconstruction of [Goldewijk et al. \(2017\)](#). Panels (b) and (c) zoom into the Anahuac Valley (the seat of Tenochtitlán and the Mexica state) and the Atemajac Valley (the 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.9: Colonial population estimates for Mexico City and Guadalajara



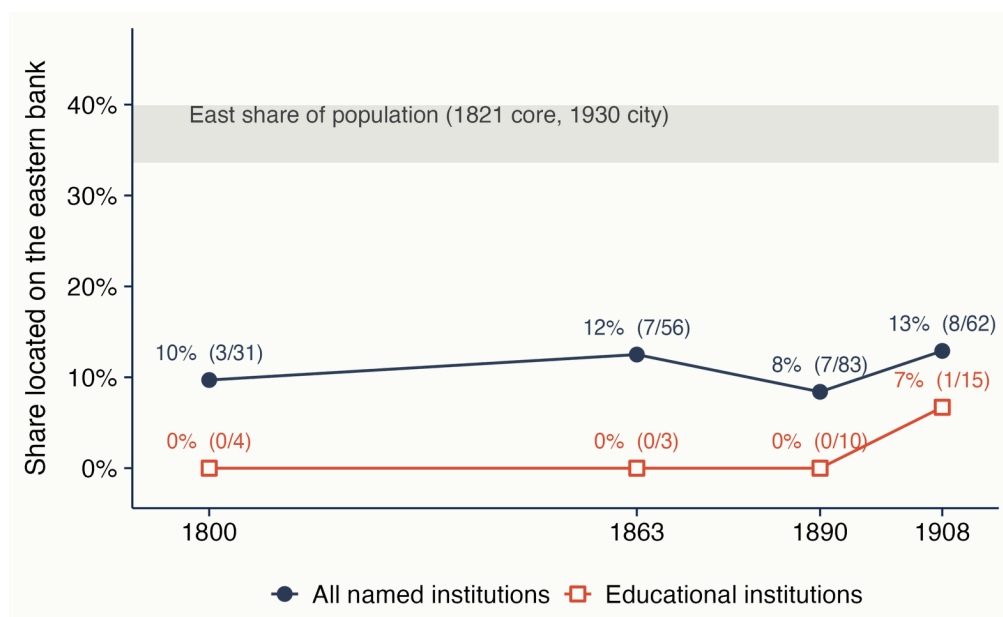
Notes: Population estimates for Mexico City and Guadalajara from the 16th to 18th centuries, based on the historical urban-population reconstructions of [Buringh \(2018\)](#). The persistent population gap between the two cities — Mexico City consistently 5–10 times larger than Guadalajara through the 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.

Figure A.10: 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.

Figure A.11: 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 BD design; 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.4 reports the sources of the four maps, the location method for every entry, and the counts by category.

Figure A.12: Cartographic sources for the count of named institutions, 1800–1908



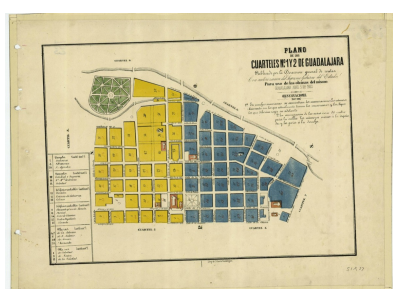
(a) 1800 facsimile plan



(b) 1890 *Plano General*



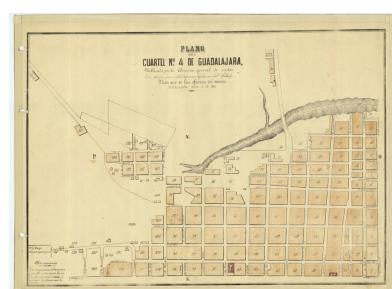
(c) 1908 *Novísimo Plano General*



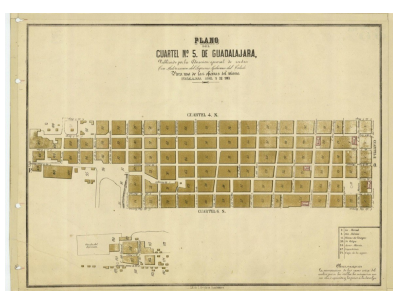
(d) 1863, *cuarteles* 1 and 2



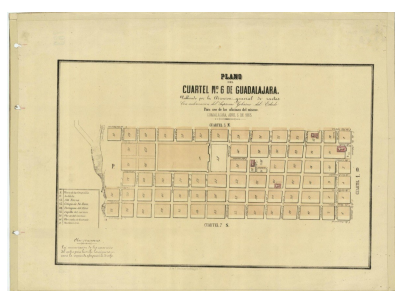
(e) 1863, *cuartel* 3



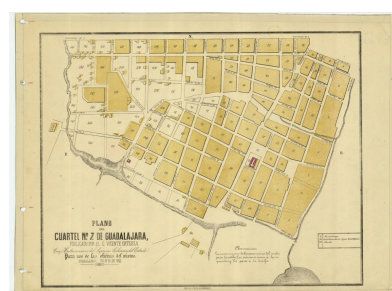
(f) 1863, *cuartel* 4



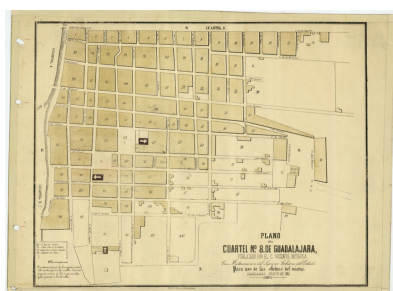
(g) 1863, *cuartel* 5



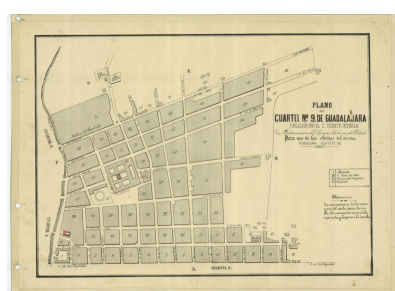
(h) 1863, *cuartel* 6



(i) 1863, *cuartel* 7



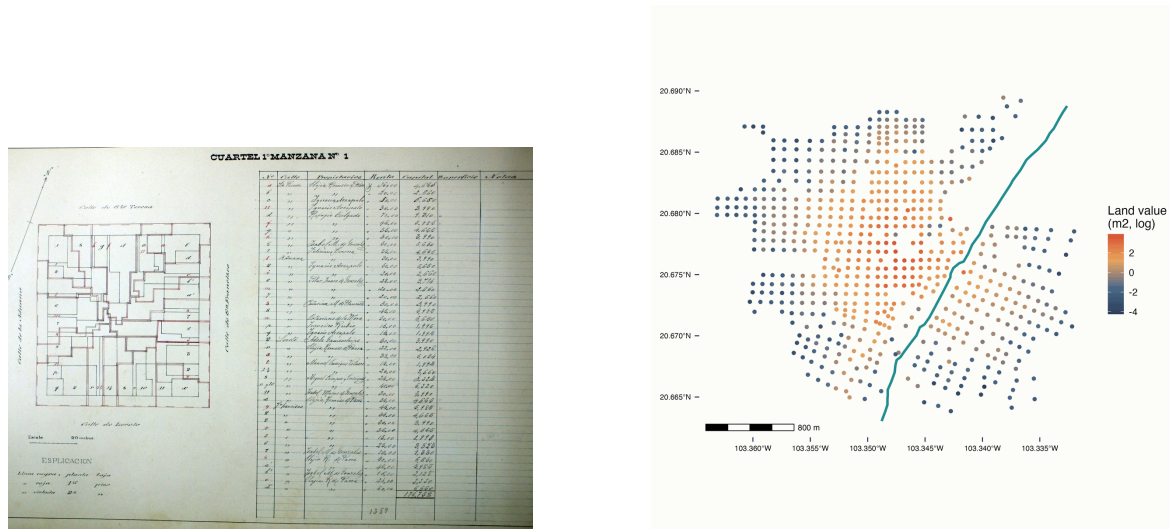
(j) 1863, *cuartel* 8 (Analco)



(k) 1863, *cuartel* 9

Notes: The four waves of Table A.2 and Figure A.11. 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.

Figure A.13: Historical land values in 1882

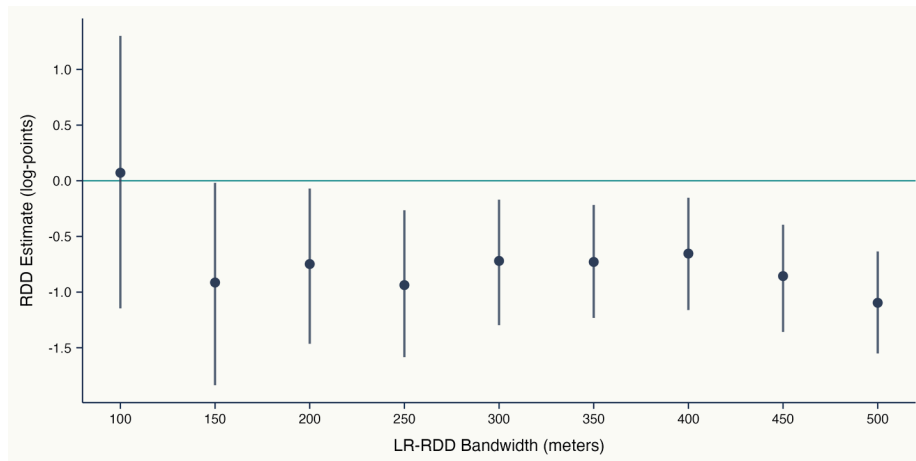


(a) Property façade from the 1882 cadaster

(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 SJD 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.

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

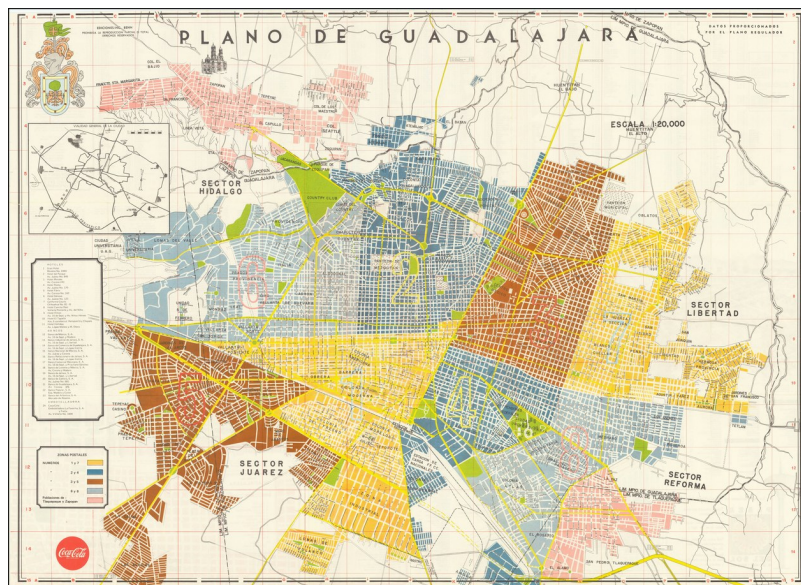


Notes: Each point reports the estimated discontinuity in log land values per square meter across the SJD creek in 1882, using the Local Randomization RDD (LR-RDD) of Cattaneo et al. (2015), which treats observations within a small window around the boundary as if assigned at random. Estimates are reported 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.

Figure A.15: Urban expansion of Guadalajara across the 20th century



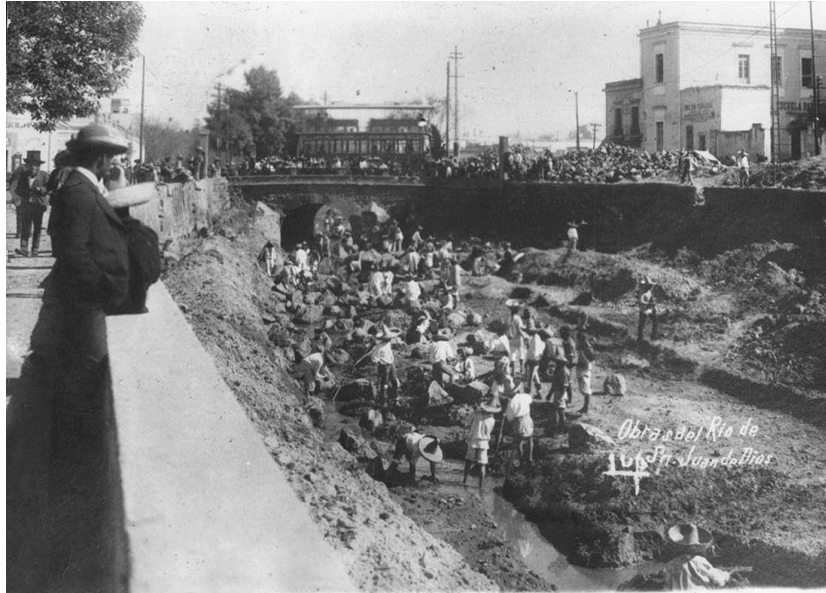
(a) Limited urban expansion by the early 20th century (1908)



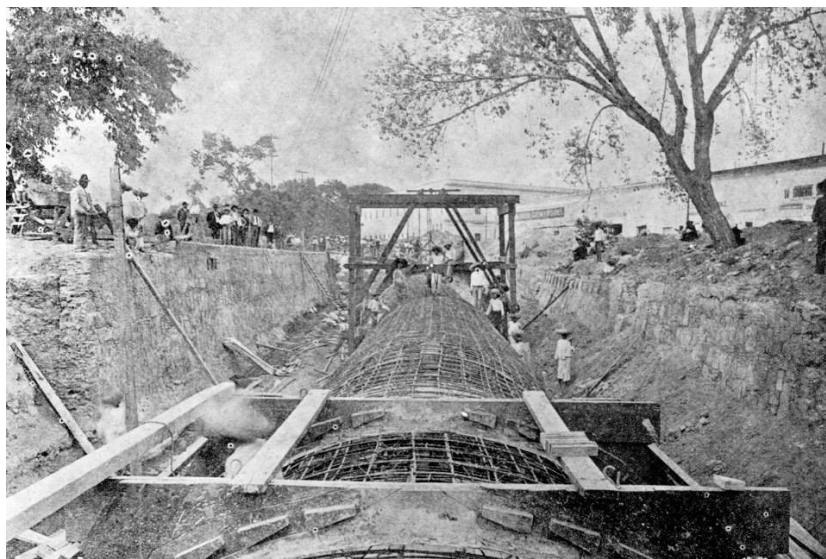
(b) Metropolitan expansion by the 1970s

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.

Figure A.16: 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.

Figure A.17: Calzada Independencia as a present-day surface-level avenue



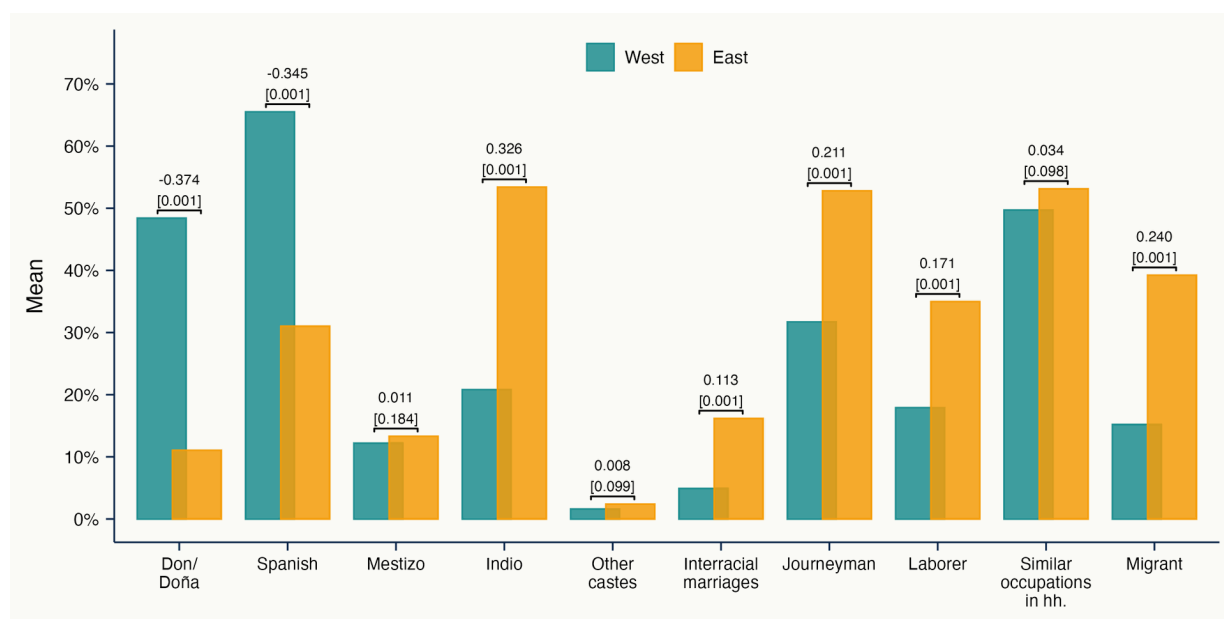
(a) Westward view of Calzada Independencia



(b) Eastward view of Calzada Independencia

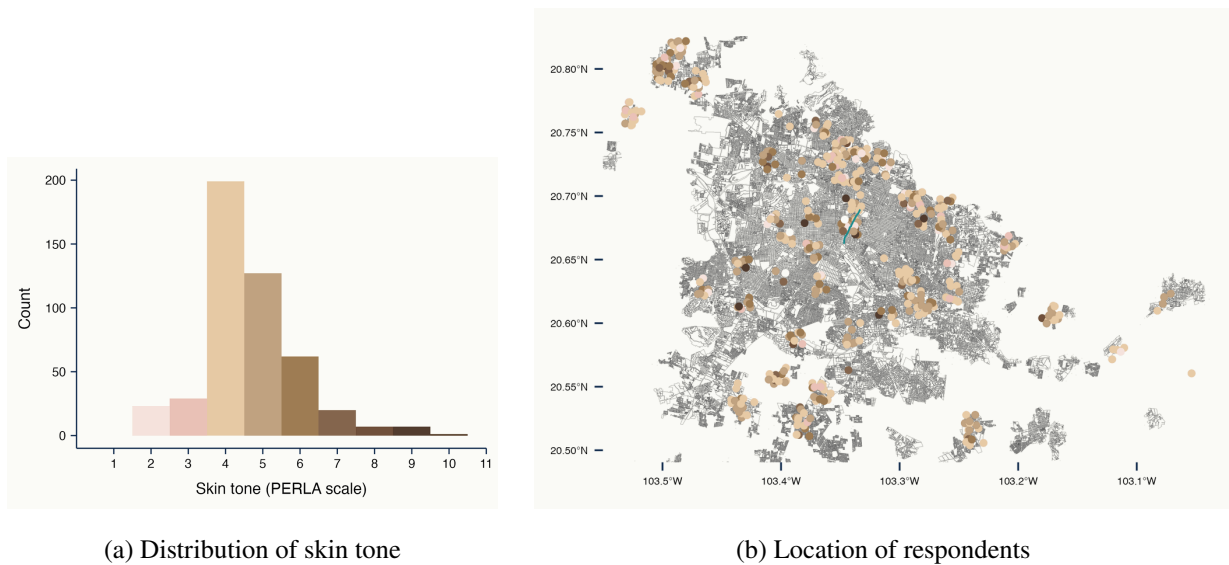
Notes: Photographs of Calzada Independencia in 2020. Panel (a) shows the view from the western (Spaniard Guadalajara and Mexicaltzingo) side; panel (b) from the eastern (Analco) side. The avenue is a surface-level urban arterial integrated into the city grid, with frequent pedestrian crossings, continuous 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 BD design estimates therefore cannot be attributed to the Calzada functioning as a contemporary physical barrier.

Figure A.18: 1821 census: socio-demographic characteristics by bank of the San Juan de Dios creek



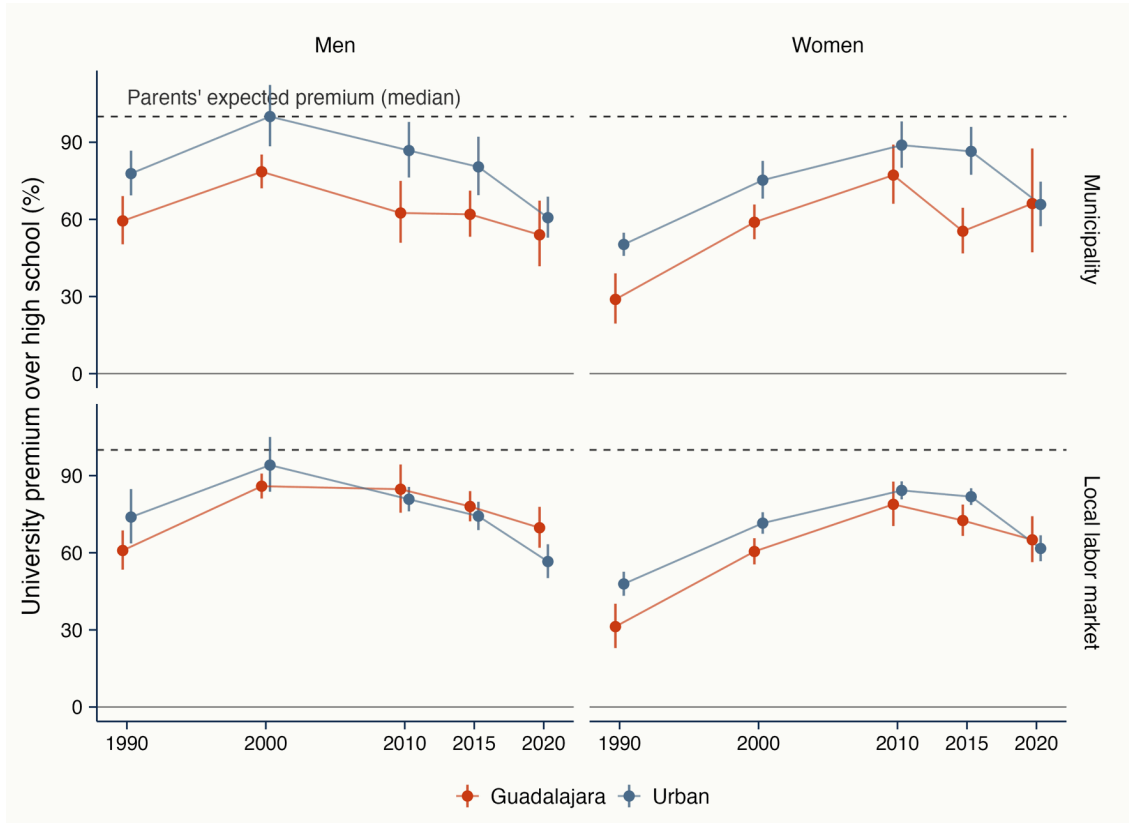
Notes: Mean values of selected socio-demographic characteristics in 1821, computed separately for residents of the eight *cuarteles* immediately west of the SJdD creek (Spaniard Guadalajara and Mexicaltzingo) and the eight *cuarteles* immediately east (Analco). Data are from the individual-level 1821 *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 BD 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.3 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 variation in colonial standing that is not reducible to ethnic composition.

Figure A.19: Skin tone across the Guadalajara metropolitan area



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.

Figure A.20: 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 4). 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.

B Persistent Human Capital Gaps

This appendix collects the supporting evidence for the Boundary Discontinuity (BD) analysis in Section 3, in the order the section draws on it. Appendix B.1 maps the schooling gradient across the metropolitan area and Appendix B.2 reports the density test. Appendix B.3 reports the full set of main estimates and robustness specifications. Appendix B.4 and Appendix B.5 report additional multi-wave outcomes and the 1930 census comparison. Appendix B.6 organizes the supply- and demand-side accounts of the gap and reports the 2015 Intercensal evidence on household resources, housing, and commuting; Appendix B.7, Appendix B.8, Appendix B.9, and Appendix B.10 then address housing-market sorting, school supply and quality, crime exposure, and street infrastructure, urban form, and businesses. Appendix B.11 estimates the discontinuity point by point along the boundary. Appendix B.12 addresses residential turnover and selective immigration. Appendix B.13 reports the donut estimates and Appendix B.14 the placebo-boundary exercise. Appendix B.15 details the design behind the 1992-explosions evidence.

B.1 The east–west pattern beyond the municipality

Figure B.1 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 of the historical core, into Zapopan, hold the highest mean schooling, and blocks east of it, into Tonalá and San Pedro Tlaquepaque, the lowest.

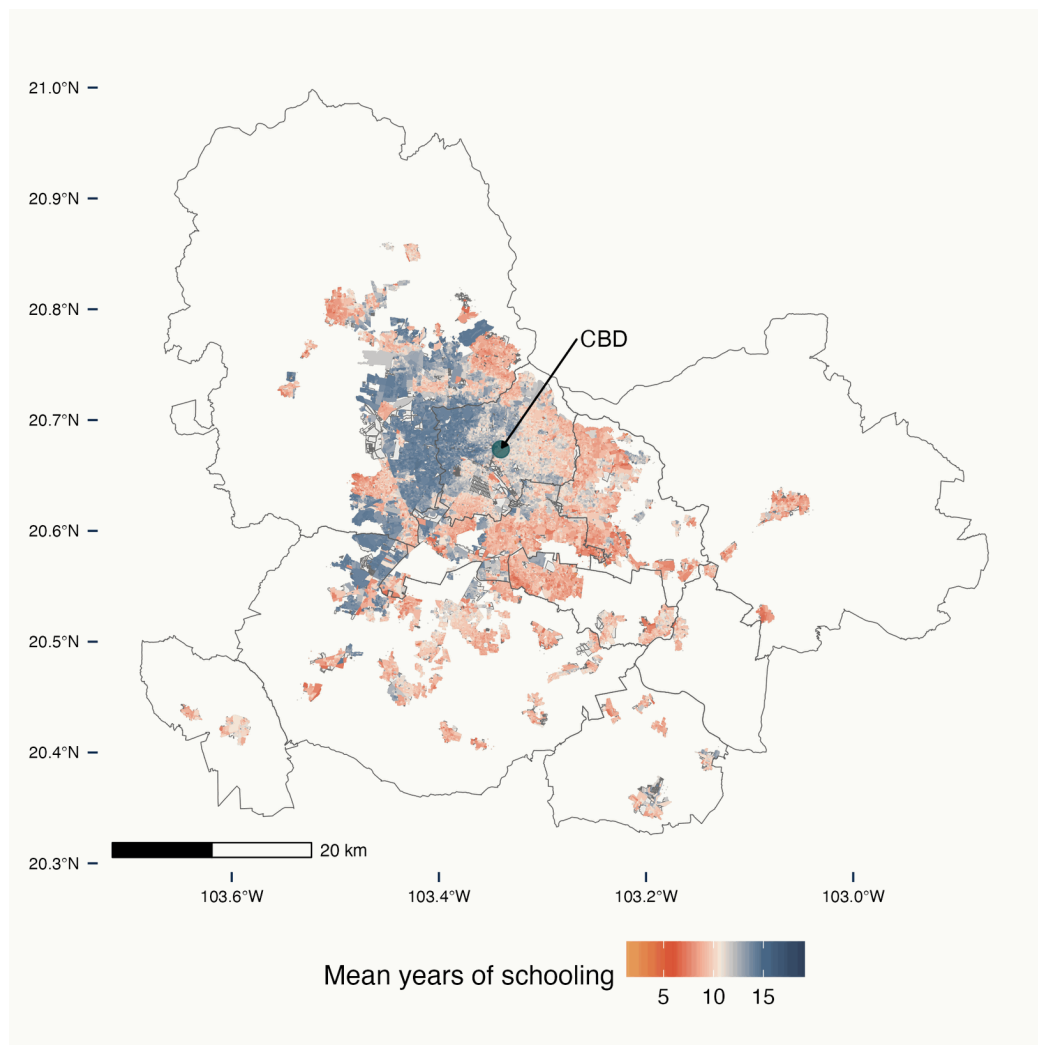
B.2 Density at the boundary

A first concern in BD designs is whether the density of observations is itself discontinuous at the boundary (McCrary, 2008; Cattaneo et al., 2018). Figure B.2 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 sides 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.3 Main estimates and full robustness battery

Table B.1 reports the full set of BD 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

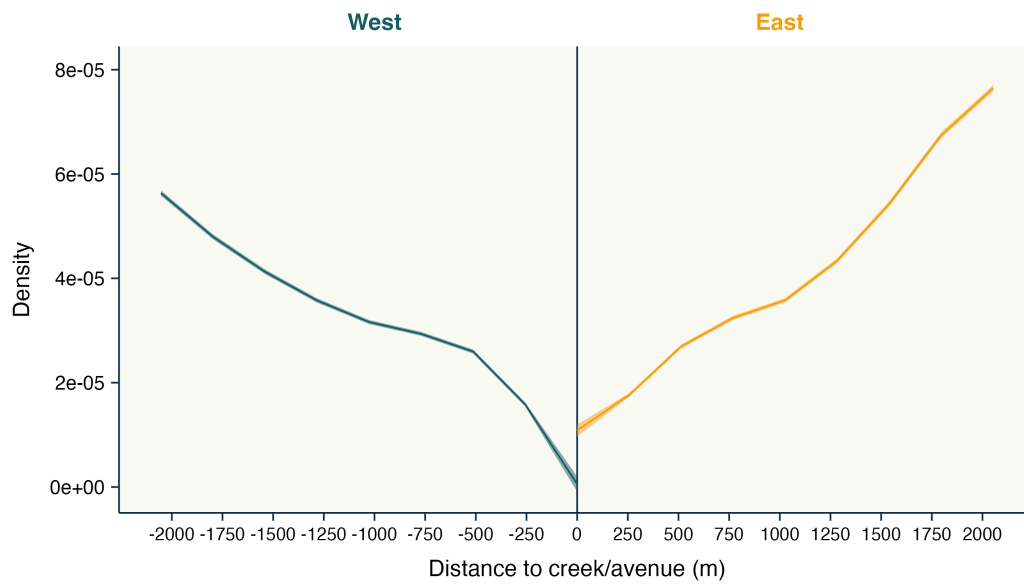
Figure B.1: Mean years of schooling by census block, Guadalajara metropolitan area (2020)



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.

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, with -0.17 in the preferred specification without covariates and -0.14 once covariates are added. School-attendance estimates are noisier but consistently negative (-0.08 to -0.14 SD). The Calonico–Cattaneo–Titiunik nearest-neighbor (NN) variance estimator is used in the main specifications; the HC3 row reports the most conservative heteroskedasticity-consistent alternative.

Figure B.2: 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 sides, and the test has limited power where the density approaches zero.

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

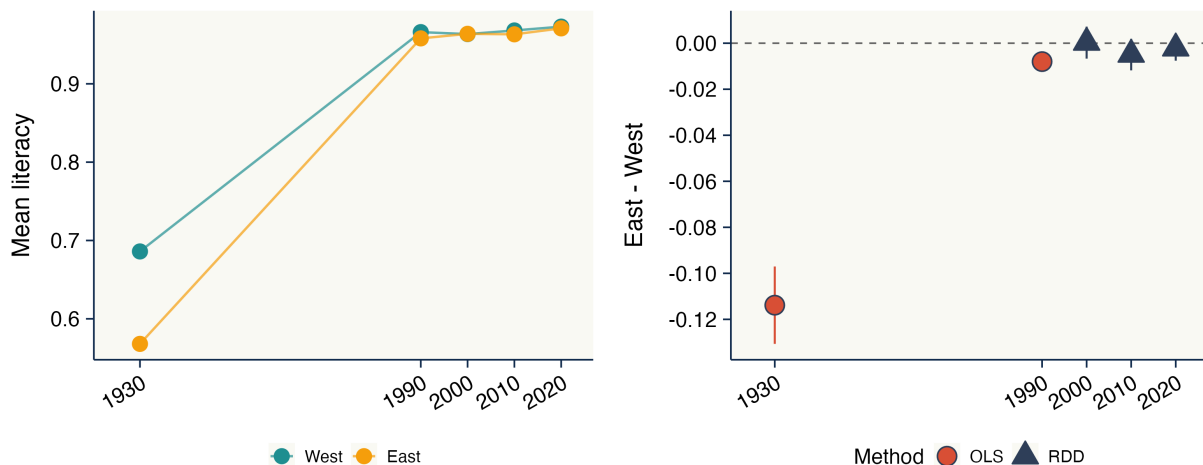
$\hat{\tau}$	SE	95% CI	p -val	Poly	h	N_h	b	N_b	Kernel	VCE	Clean ctrls	Unclean ctrls	Specification
<i>Panel A. Years of schooling (ages 6+)</i>													
-0.791	0.156	[-1.097, -0.485]	0.000	1	570.750	25011	1489.073	107104	Uniform	NN	No	No	Main – levels
-0.165	0.033	[-0.229, -0.101]	0.000	1	570.750	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)
<i>Panel B. School attendance (ages 6–22)</i>													
-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: BD 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 first row in Panel A, in levels and without covariates, 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.4 Multi-wave evolution: additional outcomes

This subsection reports the outcomes omitted from the main-text multi-wave figure (Figure 3). Figure B.3 traces adult literacy across the boundary from 1930 to 2020: literacy is near-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.

Figure B.3: Adult literacy across the historical boundary, 1930–2020



Notes: As in Figure 3, 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 BD 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.

B.5 Historical human capital gaps in 1930

The 1930 evidence highlighted in Figure 3 draws on the broader set of variables in Table B.2, 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 BD estimates. 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.

Table B.2: Socio-demographic differences across the historical boundary, 1930 Census

Variable	West Mean	East Mean	Diff.	Std. Error	<i>q</i> -value	<i>N</i>
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.6 Robustness: supply- and demand-side determinants

Educational investment responds to what schools offer and to what households can and want to spend. The supply side is taken up in the subsections that follow: school supply and quality in Appendix B.8 and crime exposure in Appendix B.9, with residential sorting and the capitalization of the boundary into land values in Appendix B.7. This subsection collects the remaining demand-side accounts, using the 2015 *Encuesta Intercensal*, which records household resources and housing that the full-count census does not. Table B.3 reports BD estimates at the boundary for each of these measures. They cover several of the determinants discussed in Section 3.5 — household resources, credit access, and parental human capital — but not all of them: wealth, liquidity, collateral, and the concentration of landownership are not measured here (Galor, 2011).

Table B.3: Household resources, housing, and commuting at the boundary (2015 Intercensal Survey)

	Clean controls			+ dirty controls		
	$\hat{\tau}$	SE	N_h	$\hat{\tau}$	SE	N_h
<i>A. Human capital</i>						
Years of schooling	-0.305***	(0.097)	3,077	-0.213**	(0.101)	2,481
Attends school in Guadalajara	-0.037	(0.032)	1,038			
<i>B. Household resources</i>						
Household income (log)	-0.281	(0.194)	832	-0.102	(0.201)	592
Government transfers	-0.013	(0.073)	867	+0.009	(0.063)	1,092
Pension income	-0.067	(0.092)	832	+0.001	(0.057)	1,632
International remittances	+0.002	(0.027)	2,745	-0.007	(0.035)	1,820
Domestic remittances	-0.005	(0.051)	1,551	-0.016	(0.046)	2,029
<i>C. Housing tenure</i>						
Owner-occupied	+0.141	(0.102)	837	+0.247***	(0.068)	1,345
Rented	-0.170	(0.116)	561	-0.304***	(0.061)	2,129
Owned by relatives, or lent	-0.098	(0.067)	786	-0.056	(0.055)	1,319
Title in the respondent's name	+0.002	(0.060)	2,036	+0.158	(0.124)	512
<i>D. How the dwelling was acquired</i>						
Bought ready-built	+0.271*	(0.148)	423	+0.503***	(0.187)	423
Had it built	-0.044	(0.085)	517	-0.040	(0.069)	603
Built by the household	-0.035	(0.077)	470	-0.039	(0.093)	342
Inherited	-0.253*	(0.134)	602	-0.401**	(0.157)	427
<i>E. How the dwelling was financed</i>						
Own resources	+0.143*	(0.082)	686	+0.169***	(0.050)	2,136
Loan from family or friends	+0.055*	(0.033)	1,359	+0.021	(0.019)	2,365
Bank or other institutional credit	-0.014	(0.027)	354	+0.086***	(0.023)	1,151
Public housing credit	+0.007	(0.025)	533	-0.011	(0.025)	1,161
Fully paid	+0.760**	(0.331)	111	+0.257	(0.197)	154
Still paying	-0.760**	(0.331)	111	-0.258	(0.197)	154
<i>F. Commute to school</i>						
Up to 15 minutes	-0.227*	(0.126)	591			
16 to 30 minutes	+0.083	(0.068)	1,229			
31 to 60 minutes	+0.059	(0.060)	536			
More than 60 minutes	+0.003	(0.016)	734			
Public transport	+0.174**	(0.081)	1,186			
Private transport	+0.104	(0.100)	255			
<i>G. Commute to work (municipality-wide comparison, not a BD estimate)</i>						
Up to 15 minutes	-0.026	(0.017)				
16 to 30 minutes	-0.037**	(0.016)				
31 to 60 minutes	+0.037**	(0.015)				
More than 60 minutes	+0.058***	(0.009)				
Walking	-0.017	(0.014)				
Public transport	+0.079***	(0.018)				
Private transport	-0.059***	(0.018)				
Works in Guadalajara	+0.051***	(0.012)				

Notes: BD estimates at the historical boundary using the 2015 *Encuesta Intercensal* for the municipality of Guadalajara, produced in the INEGI microdata laboratory. Local linear polynomial, uniform kernel, nearest-neighbor variance, survey weights, and an MSE-optimal bandwidth selected in each row (Calónico et al., 2014). *Clean controls* are age, sex, longitude, latitude, and boundary-segment fixed effects for person-level outcomes, and longitude, latitude, and boundary-segment fixed effects for household-level outcomes. *Dirty controls* add log household income per capita for person-level outcomes and, for household-level outcomes, log distance to the central business district, the sex and age of the household head, parental schooling, and log household income per capita; income is omitted from the controls in the row where it is the dependent variable. Years of schooling are in standard deviations and household income in log points; all other outcomes are in percentage points. In Panel E the last two rows are asked only of households that used credit. Panel F covers sons and daughters aged 24 or less. Panel G is not a BD estimate: it reports the coefficient on the east-side indicator in a municipality-wide comparison with side-specific linear distance trends, estimated among workers with standard errors clustered by dwelling. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Household income. If eastern households were poorer today, the schooling gap could reflect current budgets rather than anything the assignment set in motion. Household income does not jump at the boundary, and neither do government transfers, pensions, or remittances (Table B.3).

The boundary estimate for years of schooling survives conditioning on household income per capita, falling from 0.31 to 0.21 standard deviations. The municipality-wide comparison in Table 1 gives the same answer for attendance.

Parental human capital. Parental education raises the productivity of investment in children (Cunha and Heckman, 2007), so an initial difference in human capital can reproduce itself across generations even with equal schools and perfect capital markets (Galor and Tsiddon, 1997). Parental schooling absorbs about a third of the adolescent attendance gap, which falls from 8.1 to 5.4 percentage points in Table 1. That attenuation is not a confound removed: a parent’s own schooling was itself shaped by the provision asymmetry documented in Section 2.2, so it measures how much of the gap runs through the previous generation. In the experiment, the Map-First effect is no smaller among households whose parents completed high school (Table D.9).

Credit access. With imperfect credit markets, lower initial wealth can sustain lower investment in children across generations even after provision converges (Galor and Zeira, 1993; Galor, 2011). The survey records how the dwelling was acquired and financed, whether the household is still paying for it, and whether employment carries housing credit and retirement savings. None of these jumps at the boundary, except that eastern households are more likely to own and less likely to rent, which is not the pattern of a more constrained side (Table B.3). These measures capture the use of housing finance and formal access through employment, not liquidity or collateral, and they describe households today. They do not speak to whether credit constraints operated when the schooling stock was formed.

Access to schools and to jobs. Schools have no residential catchments here, and the commute to school differs little across the boundary: the shares reaching school in each travel-time band are statistically indistinguishable, and only public transport shows a difference (Table B.3). The commute to work is another matter. In the municipality-wide comparison, eastern workers are more likely to travel more than an hour and to use public transport. Access to jobs therefore differs across the two sides, while access to schools does not.

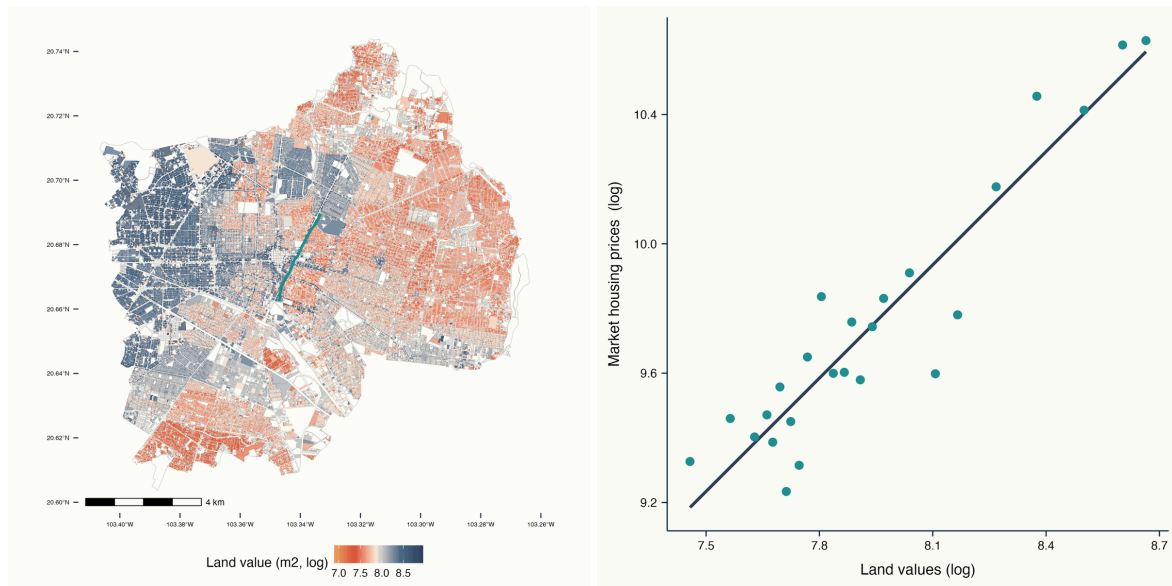
B.7 Housing-market sorting and capitalization

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.4a maps parcel-level cadastral land values, documenting a clear west–east gradient. Figure B.4b validates the cadastral measure against modern market housing prices at the neighborhood (*colonia*) 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.4: 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 (2017). 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 (2017); IIEG web-scraped housing listings.

Land-value differences by land use. Table B.4 estimates the discontinuity in parcel-level land values separately by the land use of the parcel's block. Blocks are classified with the 2020 census and the national business directory (DENUE). An establishment serves the neighborhood when it employs at most ten people, is not located in a market, shopping center, or commercial building, and belongs to a listed set of everyday activities such as grocery stores, tortillerías, pharmacies, small restaurants, beauty salons, and repair shops. A block is residential if it has inhabited dwellings and only neighborhood-serving establishments, non-residential if it has no inhabited dwellings or more than half of its establishments do not serve the neighborhood, and mixed otherwise. Land values are lower east of the boundary for non-residential parcels (-0.30 log points) and mixed parcels (-0.25), while the estimate for residential parcels (-0.21) is imprecise and not statistically significant. The differences between classes are not statistically significant.

Robustness of the boundary effect to land-value controls. Table B.5 reports BD estimates after adding flexible controls for block-level land values. The preferred specification adds land

Table B.4: BD estimates on parcel land values by block land use

Block class	$\hat{\tau}$	SE	95% CI	p -val	h	N_h
Residential	-0.208	0.541	[-0.883, 1.238]	0.743	487.8	495
Mixed	-0.249	0.207	[-0.887, -0.075]	0.020	694.8	3,471
Non-residential	-0.302	0.065	[-0.487, -0.232]	0.000	811.7	7,702
<i>Differences</i>						
Mixed – Residential	-0.040	0.579	[-1.794, 0.477]	0.256		
Non-residential – Residential	-0.094	0.545	[-1.605, 0.531]	0.325		
Non-residential – Mixed	-0.054	0.217	[-0.304, 0.547]	0.575		

Notes: BD estimates of the boundary effect on the log of cadastral land value per m² at the parcel level, estimated jointly by block land-use class with class-specific MSE-optimal bandwidths (Calónico et al., 2014). Each parcel takes the class of the census block that contains its centroid. Covariates for efficiency: longitude, latitude, and boundary-segment dummies. $\hat{\tau}$ is the conventional estimate; standard errors, confidence intervals, and p -values are robust bias-corrected, clustered by census block. h is the main bandwidth (meters) and N_h the number of parcels within it. Differences between classes use the bias-corrected estimates and treat the classes as independent subsamples.

values to the clean controls, giving -0.40 standardized block-level units, or about 0.80 years of schooling. Across the land-value specifications the estimate ranges from -0.40 to -0.48 in block-level standard deviations, that is from 0.80 to 0.97 years, close to the estimate obtained on individual data; land values absorb part of the cross-boundary differential without changing the magnitude the individual-level design reports.

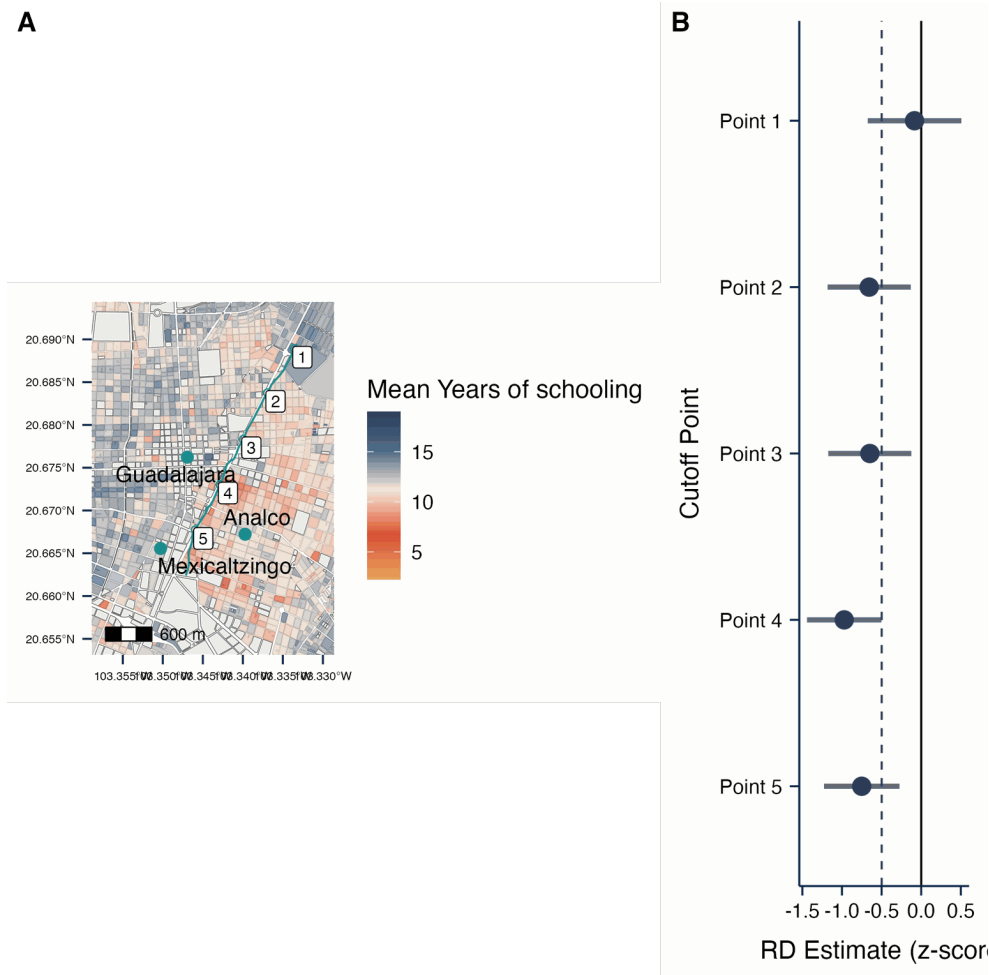
Table B.5: BD estimates controlling for land values (block level, 2020 Census)

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	-0.643	0.164	[-0.965, -0.321]	0.000	1	691.2	1645.9	463
With clean controls (Poly 2)	-0.733	0.167	[-1.060, -0.406]	0.000	2	1542.1	2767.6	1,397
+ Land value control	-0.402	0.143	[-0.682, -0.122]	0.005	1	804.8	1724.2	570
+ Land value control (Poly 2)	-0.483	0.145	[-0.766, -0.199]	0.001	2	1779.5	2866.4	1,687

Notes: BD 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 BD estimate without land-value controls. “Clean controls” include block longitude, latitude, and boundary-segment dummies. “Land value control” adds the log of cadastral land values (pesos per m²). 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 Calónico 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 0.80 and 0.97 years once block-level standardized units are converted at the block standard deviation of 2.00.

Location-based estimates with land-value controls. Figure B.5 extends the location-based BD estimator of Cattaneo et al. (2025) to the land-value-controlled specification. The pointwise estimates remain negative and concentrated near the historical core.

Figure B.5: Location-based BD estimates with block-level land-value controls



Notes: Location-based BD 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.5.

B.8 School supply and quality

School supply. Table B.6 reports BD estimates of the boundary effect on the number of schools per block, by sector, level, and shift, from 2018–2019 administrative records, together with the public share of schools and the number of students per teacher. The sample includes every block in Guadalajara, so blocks without a school enter with zeros. None of the nine outcomes shows a significant discontinuity, and all q -values equal one.

Table B.6: BD estimates on school supply (block level, 2018–2019)

Outcome	$\hat{\tau}$	SE	95% CI	p -val	q -val	Poly	h	b	N	Mean
Total schools	0.019	0.062	[-0.102, 0.140]	0.754	1.000	1	2149.2	3190.1	5217	0.111
Total schools $\times$ shifts	0.010	0.070	[-0.128, 0.147]	0.890	1.000	1	2178.4	3180.3	5188	0.121
Public schools	0.014	0.042	[-0.068, 0.096]	0.743	1.000	1	2109.3	3257.3	5407	0.064
Private schools	-0.003	0.055	[-0.111, 0.105]	0.956	1.000	1	1932.6	2856.4	4401	0.057
Primary schools	0.004	0.025	[-0.044, 0.052]	0.874	1.000	1	2151.0	3269.9	5445	0.041
Secondary schools	-0.006	0.012	[-0.029, 0.017]	0.606	1.000	1	2066.0	3378.2	5754	0.015
High schools	-0.051	0.032	[-0.113, 0.012]	0.111	1.000	1	3330.2	4507.3	8775	0.015
Public share	0.255	0.240	[-0.215, 0.725]	0.288	1.000	1	1366.1	2296.8	227	0.447
Students per teacher	-0.255	4.310	[-8.703, 8.193]	0.953	1.000	1	2420.7	3877.0	466	18.5

Notes: BD estimates of the boundary effect on educational infrastructure, using administrative school records for 2018–2019 geocoded to the 13,791 census blocks of Guadalajara. Dependent variables, in levels: total schools per block; school shifts (*turnos*); public, private, primary, secondary, and high-school counts; the public share of schools; and students per teacher. Blocks without a school enter with zeros in the count outcomes; the public share and students per teacher are defined only for blocks with at least one school. Inference follows Calonico et al. (2014); specifications include boundary-segment fixed effects and geographic coordinates. N counts blocks within bandwidth b ; Mean is the sample mean of the dependent variable. The q -value column reports sharpened FDR-adjusted p -values across the nine outcomes (Anderson, 2008).

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.6 and B.8) and a BD design within 1.5 km of the boundary (Figures B.7 and B.9). All four exercises show no systematic quality differential across the boundary. School access and quality are not the channel.

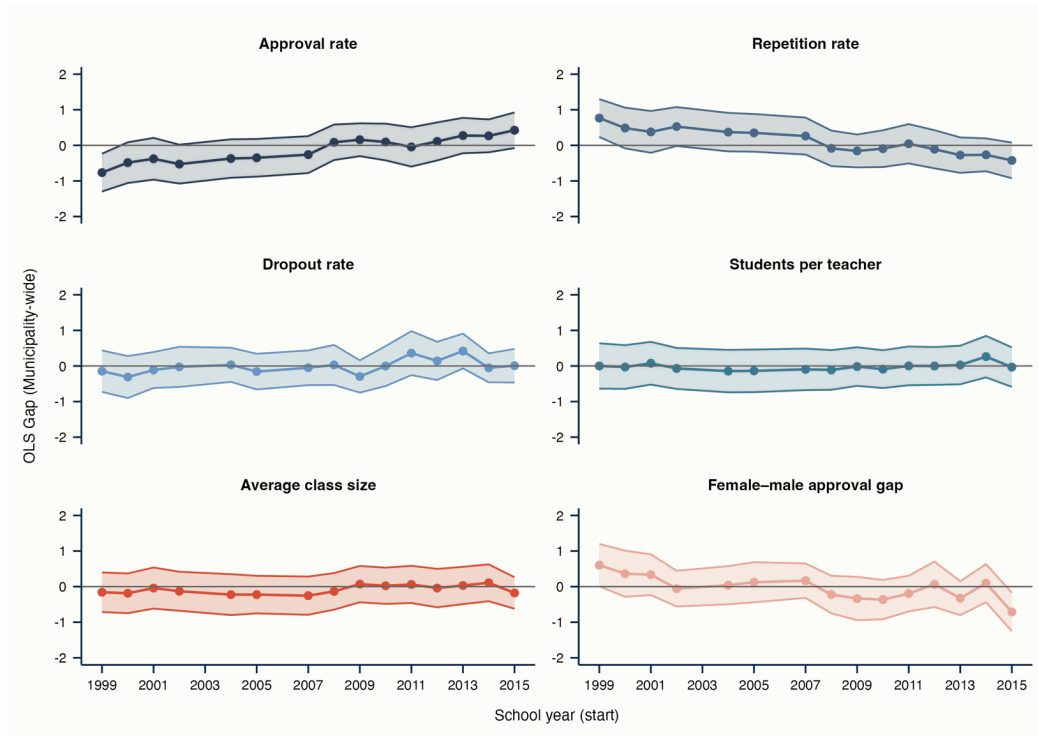
B.9 Crime exposure

Figure B.10 maps reported crimes (*carpetas de investigación*) in central Guadalajara (2017–2024). Table B.7 formalizes the comparison: none of the ten estimates is significant after FDR adjustment, and the largest unadjusted estimates (violence against women, $p = 0.055$ and $p = 0.110$) have q -values of one. The boundary does not coincide with a discontinuity in crime exposure.

B.10 Street infrastructure, urban form, and businesses

Table B.8 reports BD estimates for eleven features of the street that the 2020 National Housing Inventory records on each block front. Paved streets, sidewalks, wheelchair ramps, and mobile street stalls are more frequent on the eastern side, and no feature is less frequent there. Table B.9 reports block-level measures of urban form and the area of parcels. Residents and occupied dwellings per block and parcel area do not differ at the boundary. Tables B.10 and B.11 use the establishments of the December 2019 business directory. Establishments east of the boundary

Figure B.6: Municipality-wide OLS estimates of school quality (Formato 911, secondary)



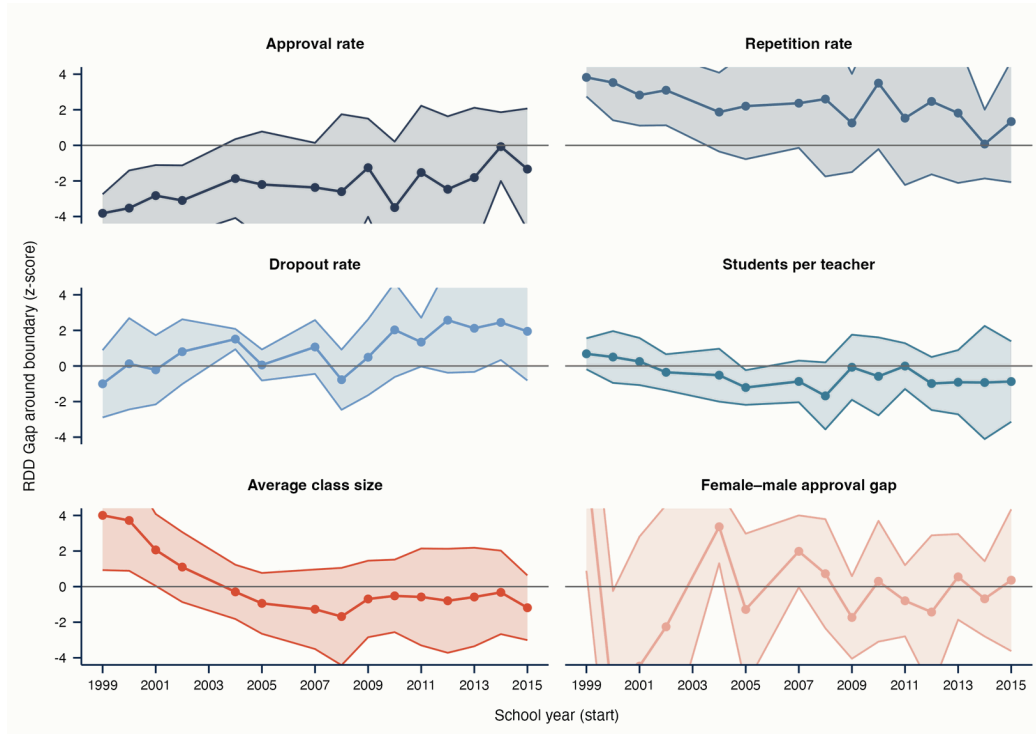
Notes: East-minus-west gaps in school-level operational indicators for secondary schools across Guadalajara, using *Formato 911* data (1998–2015). Each panel reports a separate outcome (approval, repetition, dropout, class size, student–teacher ratio, female-minus-male approval gap), standardized within 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.

Table B.7: BD estimates on crime opportunities (block level, 2017–2024)

Outcome	$\hat{\tau}$	SE	95% CI	p -val	q -val	Poly	h	b	N
<i>Panel A. Total crime</i>									
Log density	0.145	0.247	[-0.340, 0.630]	0.557	1.000	1	737.3	1164.1	1,246
Log rate	0.055	0.239	[-0.413, 0.523]	0.818	1.000	1	807.8	1184.5	1,277
<i>Panel B. By crime type</i>									
Property crimes (log density)	0.053	0.200	[-0.340, 0.446]	0.791	1.000	1	723.1	1137.5	1,211
Property crimes (log rate)	-0.036	0.203	[-0.434, 0.361]	0.858	1.000	1	797.0	1158.9	1,238
Violent crimes (log density)	0.038	0.072	[-0.103, 0.178]	0.601	1.000	1	821.9	1257.2	1,379
Violent crimes (log rate)	0.017	0.079	[-0.138, 0.171]	0.831	1.000	1	865.7	1282.7	1,408
Homicide (log density)	0.017	0.018	[-0.019, 0.053]	0.361	1.000	1	783.5	1353.4	1,506
Homicide (log rate)	0.022	0.016	[-0.009, 0.053]	0.164	1.000	1	1008.3	1522.6	1,749
Violence Against Women (log density)	0.165	0.086	[-0.004, 0.333]	0.055	1.000	1	839.6	1274.5	1,400
Violence Against Women (log rate)	0.117	0.073	[-0.026, 0.260]	0.110	1.000	1	822.8	1256.7	1,378

Notes: BD 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 km²; 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, femicide. Inference follows [Calonico et al. \(2014\)](#); specifications include boundary-segment fixed effects and geographic coordinates. N counts blocks within bandwidth b . The q -value column reports sharpened FDR-adjusted p -values across the ten outcomes ([Anderson, 2008](#)). None is significant after correction.

Figure B.7: BD estimates of school quality (Formato 911, secondary)



Notes: BD 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.

are smaller, with fewer establishments of more than thirty employees and fewer establishments per resident, and the sector shares do not differ after the false-discovery-rate adjustment.

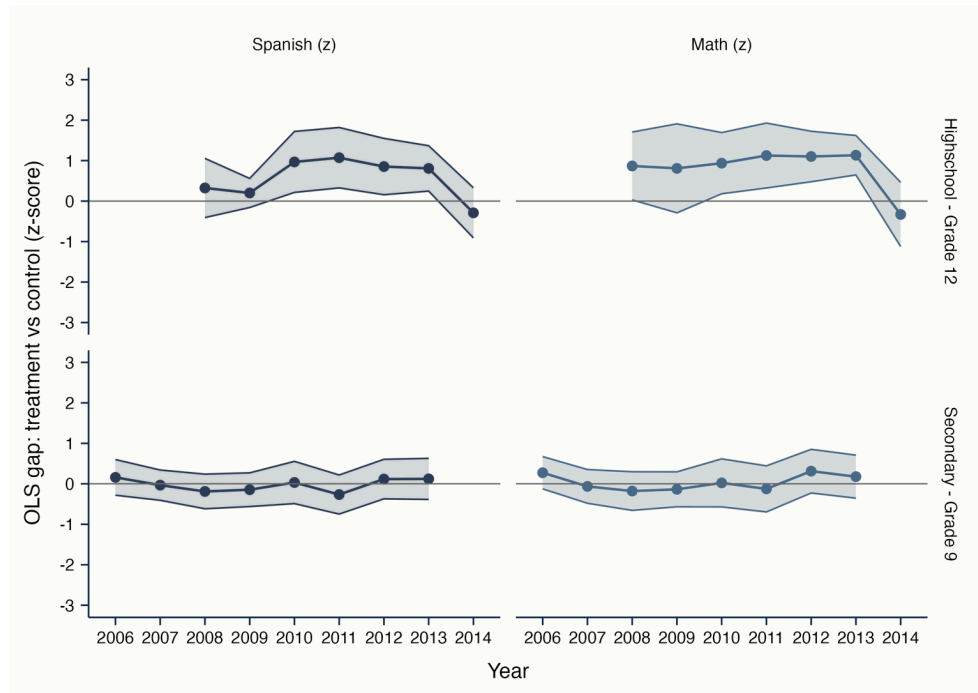
B.11 Estimates along the boundary

Because the historical boundary is irregular, Figure 4 estimates the discontinuity at each point along it. Table B.12 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.12 Residential mobility and selective in-migration

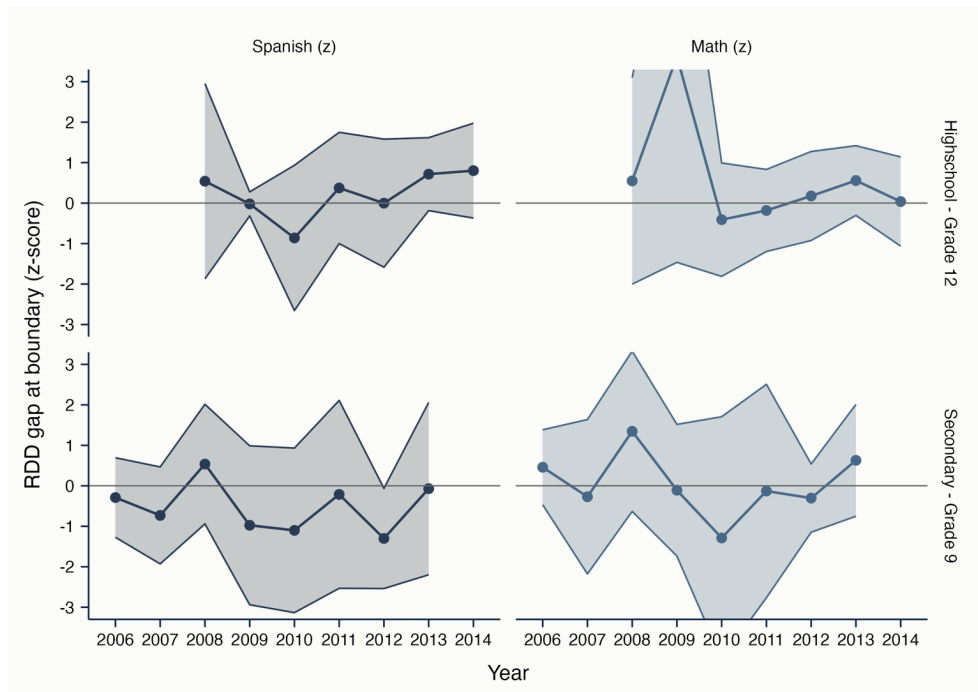
According to the 2020 census, 91.5% of Guadalajara residents lived in Jalisco five years earlier, and re-estimating the BD design excluding inter-state migrants leaves the results essentially unchanged (last row of Panel A, Table B.1). To approximate within-city mobility, I match adults aged 34+ in 2020 to plausibly the same individuals in the same block in 2010, using nearest-

Figure B.8: 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.6, and the gaps fluctuate around zero throughout.

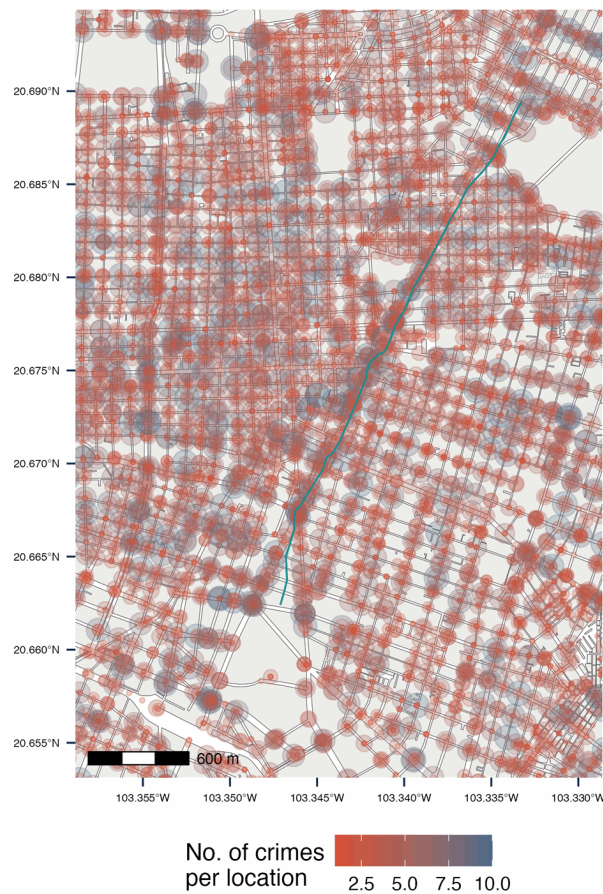
Figure B.9: BD estimates of standardized test performance (ENLACE)



Notes: BD 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.7, and the estimates fluctuate around zero throughout.

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.13 reports the boundary

Figure B.10: 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.7.

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 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.2), 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 the mobility evidence does not show the excluded side to be 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. It does not rule out selective migration more generally. What matters for the schooling gap is the

Table B.8: BD estimates on street infrastructure (block fronts, 2020)

Outcome	Mean within h		$\hat{\tau}$	SE	p -val	q -val	h (m)	N_h
	West	East						
Paved street (asphalt or concrete)	0.827	0.950	0.474	0.147	0.002	0.012	468	1,713
Sidewalk	0.905	0.917	0.210	0.075	0.003	0.012	816	3,339
Wheelchair ramp	0.588	0.784	0.361	0.110	0.001	0.012	584	2,174
Storm drain	0.230	0.306	0.205	0.109	0.067	0.104	702	2,759
Street lighting	0.552	0.557	0.138	0.096	0.095	0.125	760	3,037
Trees	0.716	0.658	0.029	0.108	0.749	0.599	893	3,730
Pedestrian crossing	0.841	0.845	-0.023	0.078	0.714	0.599	421	1,525
Pedestrian traffic light	0.115	0.063	0.023	0.162	0.976	0.602	660	2,552
Public transport stop	0.071	0.087	-0.082	0.111	0.412	0.347	919	3,869
Street stall (semi-fixed)	0.063	0.062	0.123	0.146	0.239	0.212	726	2,881
Street stall (mobile)	0.061	0.078	0.279	0.129	0.011	0.024	705	2,777

Notes: BD estimates of the boundary effect on street infrastructure, using the block fronts of the 2020 National Housing Inventory (*Inventario Nacional de Viviendas*) within 3 km of the former creek. Each row is a separate estimate, east minus west, from a local linear regression with an MSE-optimal bandwidth h , separate linear fits on each side, and longitude, latitude, and boundary-segment fixed effects as covariates; standard errors and p -values are robust bias-corrected (Calónico et al., 2014). Each outcome equals one if the inventory records the feature on the front and is missing when it does not apply or is not specified; paved streets are asphalt or concrete. The West and East columns are raw shares within h of the boundary; $\hat{\tau}$ and its standard error are in standard deviations of the outcome. The side and the boundary segment are those of the front's census block. N_h is the number of fronts within h . q -values are sharpened false-discovery-rate values across the eleven outcomes (Anderson, 2008).

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.7 observes it, and the two sit on either side of it.

B.13 Donut estimates

B.14 Placebo boundaries and recentered specification

I randomly shift and rotate the boundary 1,000 times within the municipality and re-estimate the BD design at each placebo location. Three examples are shown in Figure B.11; each preserves the

Table B.9: BD estimates on urban form (blocks and parcels)

	Mean within h							
Outcome	West	East	$\hat{\tau}$	SE	p -val	q -val	h (m)	N_h
<i>Panel A. Blocks (2020 Census, 2017 cadastre)</i>								
Residents per block	75.3	92.1	2.273	8.911	0.978	0.432	969	735
Occupied dwellings per block	23.2	27.2	-0.740	2.966	0.601	0.430	922	682
Occupied dwellings per hectare	66.3	48.0	-12.878	13.145	0.211	0.178	1,620	1,487
Log median parcel area (m ²)	5.170	5.126	0.010	0.095	0.818	0.432	1,483	1,272
Share of parcels in condominium	0.009	0.007	-0.015	0.015	0.179	0.176	2,001	2,035
<i>Panel B. Parcels (2017 cadastre; SEs clustered by block)</i>								
Log parcel area (m ²), parcel level	5.165	5.130	0.007	0.061	0.927	–	1,456	29,102

Notes: BD estimates of the boundary effect on urban form. Panel A uses the 12,077 populated census blocks of Guadalajara, with the 2020 Census and the 2017 cadastre, whose 289,935 parcels are assigned to blocks by their centroid. Panel B uses parcels, with standard errors clustered by block. Each row is a separate estimate, east minus west, from a local linear regression with an MSE-optimal bandwidth h , separate linear fits on each side, and longitude, latitude, and boundary-segment fixed effects as covariates; standard errors and p -values are robust bias-corrected (Calonico et al., 2014). The West and East columns are means within h of the boundary, and N_h is the number of observations within h . q -values are sharpened false-discovery-rate values across the ten block-level outcomes estimated, of which five are shown (Anderson, 2008).

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 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 $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. 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 BD design on the residuals. The exposure measures are therefore block-level and are built from the same draws that generate the

Table B.10: BD estimates on business structure (establishments, 2019)

Outcome	Mean within h		$\hat{\tau}$	SE	p -val	q -val	h (m)	N_h
	West	East						
<i>Panel A. Establishment characteristics (establishment level)</i>								
0-5 employees	0.788	0.844	0.053	0.028	0.035	0.132	1,434	16,083
6-10 employees	0.114	0.087	-0.022	0.014	0.098	0.172	1,746	19,916
11-30 employees	0.066	0.050	-0.007	0.010	0.408	0.374	1,423	15,985
31+ employees	0.032	0.020	-0.024	0.010	0.006	0.072	1,232	14,077
Log employees (bracket midpoint)	1.444	1.344	-0.114	0.054	0.018	0.108	1,368	15,319
Registered legal name (formality proxy)	0.293	0.203	-0.040	0.030	0.282	0.320	2,075	23,932
Lists a phone	0.404	0.402	0.033	0.049	0.490	0.417	1,426	16,008
Lists an e-mail	0.256	0.179	0.048	0.035	0.060	0.158	897	10,032
Lists a website	0.122	0.068	-0.038	0.019	0.068	0.158	1,834	20,913
Inside market / plaza / commercial building	0.091	0.091	0.050	0.146	0.629	0.523	2,126	24,565
Semi-fixed unit	0.017	0.015	0.011	0.012	0.336	0.337	1,796	20,463
Neighborhood-serving activity	0.334	0.307	-0.000	0.044	0.805	0.675	1,593	17,809
<i>Panel B. Density (block level)</i>								
Block has any establishment	0.942	0.901	0.062	0.037	0.066	0.071	1,071	851
Log(1 + establishments in block)	2.089	2.050	-0.175	0.166	0.214	0.111	1,510	1,356
Establishments per 100 residents (populated blocks)	37.8	25.0	-47.301	23.581	0.018	0.059	1,745	1,677

Notes: BD estimates of the boundary effect on business structure, using the 89,893 establishments of the December 2019 *Directorio Estadístico Nacional de Unidades Económicas* (DENUE), each placed in the nearest 2020 census block within 50 meters. The running variable is the signed distance of the block to the boundary, and standard errors are clustered by block. Panel A has one observation per establishment and Panel B one per block. Each row is a separate estimate, east minus west, from a local linear regression with an MSE-optimal bandwidth h , separate linear fits on each side, and longitude, latitude, and boundary-segment fixed effects as covariates; standard errors and p -values are robust bias-corrected (Calonico et al., 2014). The West and East columns are means within h of the boundary, and N_h is the number of observations within h . q -values are sharpened false-discovery-rate values within each panel (Anderson, 2008).

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.

B.15 The 1992 explosions: design details

This subsection describes the design behind Figure 8 in Section 6. 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 side, 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

Table B.11: BD estimates on the sector mix of establishments (2019)

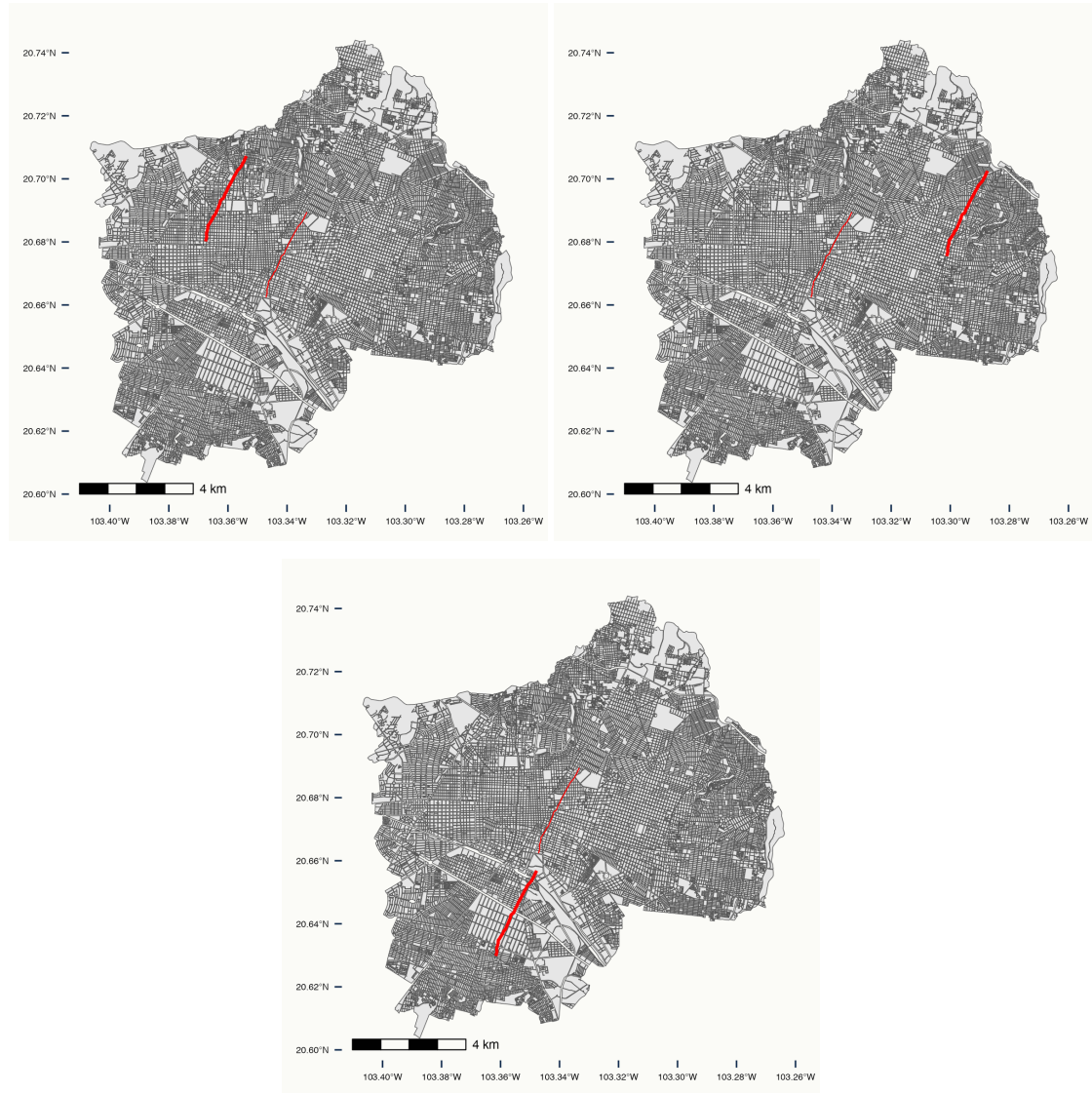
	Mean within h		$\hat{\tau}$	SE	p -val	q -val	h (m)	N_h
Outcome	West	East						
<i>SCIAN 2-digit sector shares (establishment level)</i>								
Manufacturing	0.097	0.076	-0.018	0.024	0.655	0.859	1,556	17,441
Wholesale trade	0.060	0.062	0.022	0.023	0.237	0.668	1,070	12,266
Retail trade	0.340	0.430	0.100	0.049	0.057	0.667	2,440	28,629
Information	0.003	0.001	0.001	0.002	0.545	0.859	1,632	18,265
Finance and insurance	0.017	0.010	-0.013	0.012	0.129	0.667	2,276	26,483
Real estate and rental	0.012	0.008	-0.003	0.005	0.732	0.859	2,305	26,811
Professional services	0.041	0.010	-0.009	0.010	0.644	0.859	1,245	14,141
Business support services	0.019	0.011	0.010	0.006	0.048	0.667	1,687	19,077
Education	0.025	0.016	-0.007	0.006	0.143	0.667	1,500	16,776
Health and social assistance	0.037	0.026	-0.002	0.007	0.719	0.859	1,087	12,469
Recreation	0.010	0.009	-0.001	0.003	0.665	0.859	2,554	30,482
Food and lodging	0.138	0.104	-0.006	0.028	0.967	0.936	1,741	19,801
Other services (repair, personal)	0.131	0.151	-0.028	0.026	0.198	0.667	1,382	15,493
Government	0.015	0.004	-0.013	0.009	0.092	0.667	2,244	26,040

Notes: BD estimates of the boundary effect on the share of establishments in each two-digit SCIAN sector, with the sample and the specification of Table B.10. There is one observation per establishment, and standard errors are clustered by block. q -values are sharpened false-discovery-rate values across the fourteen sectors (Anderson, 2008).

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-side AGEBs under the same specification, clustering by AGEB.

The same comparison for the composition of adults gives a difference in age of 0.88 years in 1990 (standard error 0.39), 1.35 in 2000 (0.64), -0.86 in 2010 (0.43), and 0.02 in 2020 (0.63). The change between 2000 and 2010 is -2.2 years (0.77), against a change of -1.07 years (0.43) in adult schooling. Sex, Indigenous language, and household size do not differ in any wave. In the public block tabulations, affected blocks in 2010 and 2020 have fewer residents aged 65 and older, more children per household, and fewer female-headed households than comparison blocks, while mean schooling does not differ.

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 SJD 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 BD design on census-block mean years of schooling; the resulting distribution is in Figure B.12.

Table B.12: Location-based BD estimates by boundary point

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

Notes: Location-based BD 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 4, 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 guaranteed. These are therefore pointwise contrasts rather than a simultaneous test over the family of comparisons. Stars refer to the unadjusted *g*-value of each contrast and the *q* column reports the sharpened false-discovery-rate value of (Anderson, 2008), computed across the six comparisons within each outcome. A sixth boundary point is estimated but returns no bias-corrected estimate or band, so neither the band nor the contrasts cover it. *, **, and ***

Table B.13: BD 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

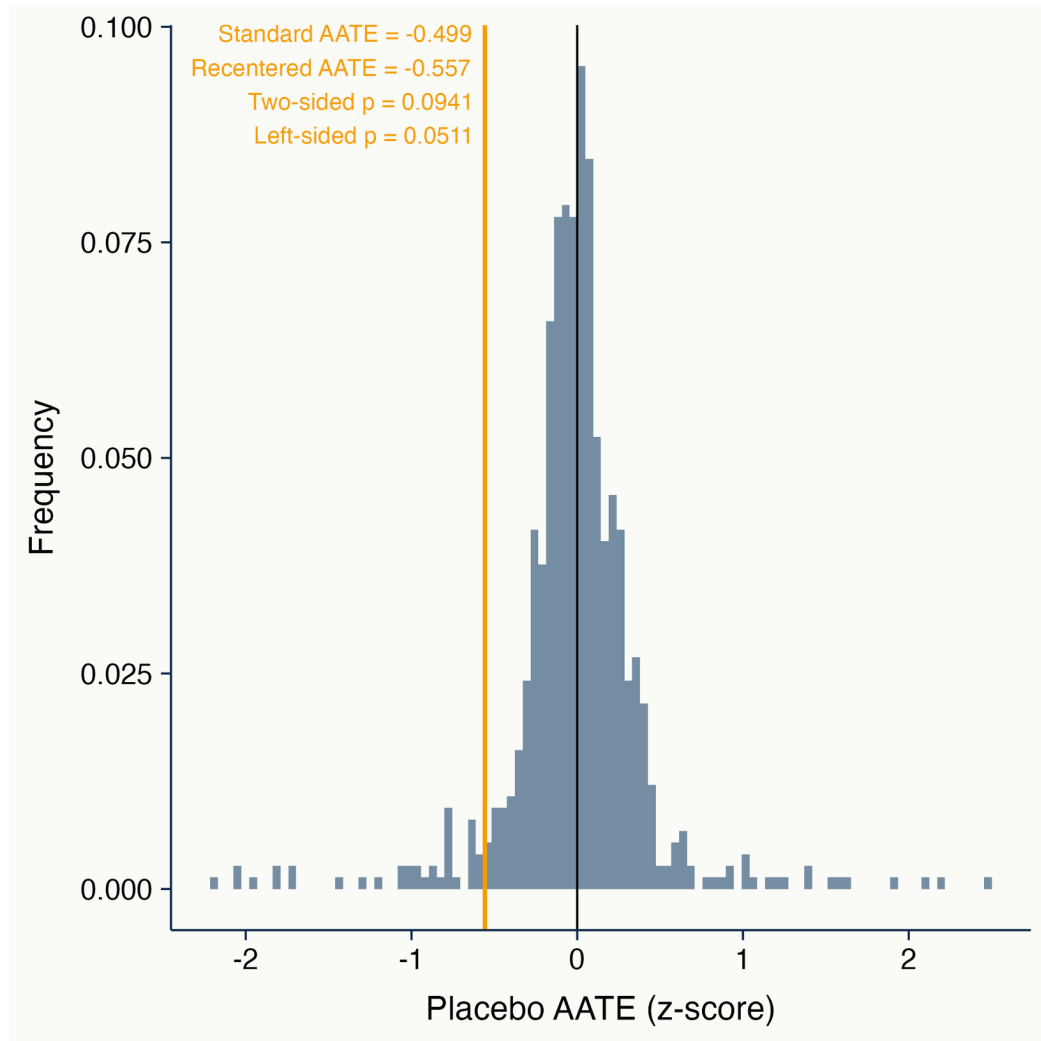
Notes: BD estimates of the boundary effect on the probability of long-term residence, individual-level 2020 census data for Guadalajara, restricted to adults aged 34+. The outcome is constructed via nearest-neighbor matching of each 2020 resident to a 2010 resident of the same block on age, sex, years 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. Because the census does not follow former residents and the indicator is not a measured turnover rate, this does not rule out selective residential sorting or selection among unobserved leavers.

Table B.14: Donut BD estimates on human capital, 2020 Census

$\hat{\tau}$	SE	95% CI	p -val	h	N_h	Donut (m)
<i>Panel A. Years of schooling (ages 6+)</i>						
-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
<i>Panel B. School attendance (ages 6–22)</i>						
-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: BD 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 side, 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.2), 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 sides are comparably populated.

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



Notes: Distribution of standardized BD 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.

C Neighborhood Social Differences: Supporting Material

This appendix collects the supporting material for Section 4.6, in the order the section draws on it. Appendix C.1 reports the additional Life Quality Survey results, and Appendix C.2 the moderated mediation analysis that earlier drafts carried in the main text. Appendix C.3 maps the parishes behind the patron-saint assignment and Appendix C.4 reports the calendar-saint placebo. Appendix C.5 documents the construction of the naming measures and reports given-name and surname distinctiveness by cohort.

C.1 Additional Life Quality Survey results

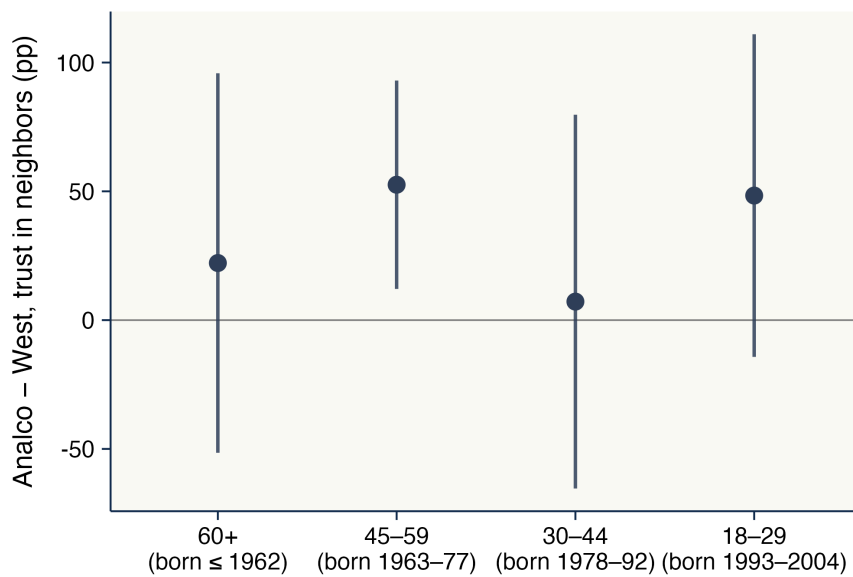
C.1.1 Trust and participation across the boundary

Table 2 reports the comparisons of schooling and interpersonal trust summarized in Section 4.1, including the relation between trust in neighbors and schooling; institutional trust and association membership are reported below. These are descriptive patterns. The survey interviews 289 Guadalajara residents in 23 census tracts, 26 of them within 1.5 kilometers of the boundary, and unconditional means are similar across sides, with 67 percent of western and 64 percent of eastern respondents reporting that they trust their neighbors and an institutional-trust index that averages 0.23 on both sides. The coefficient on the Analco indicator is the discontinuity at the boundary implied by a separate linear distance gradient on each side. All estimates use the survey's expansion factors. Panel A includes no schooling control and Panel B controls linearly for the respondent's years of schooling; the two panels give similar estimates. These estimates depend on the extrapolation.

C.1.2 The neighbor-trust gap by age bracket

Figure C.1 estimates the Analco gap in trust in neighbors separately by the Life Quality Survey's four age brackets, with the same controls as Panel B of the main specification. The point estimate is positive in every bracket, at about 48 percentage points among respondents aged 18–29, born after the early 1990s, and the four estimates are statistically indistinguishable. Each bracket has between 60 and 102 respondents, so the individual estimates are imprecise. The survey records age only in these brackets, and in a single cross-section age and birth cohort cannot be separated, so the figure documents a positive gap at every age observed rather than a cohort trend. The youngest bracket matters for the argument because it shows the gap among respondents schooled long after the provision differences of Section 2.2 had closed.

Figure C.1: 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, Panel B, estimated separately within each age bracket of the 2022 Life Quality Survey (sample sizes 60–102), with sex fixed effects, a linear control for years of schooling, side-specific distance gradients, and survey weights; vertical bars show 95 percent heteroskedasticity-robust confidence intervals. Birth cohorts follow from the 2022 survey year. The four brackets are estimated on disjoint subsamples, so the gaps are independent. A Wald test does not reject equality across them ($W = 1.45$ on three degrees of freedom, $p = 0.69$), while the hypothesis that all four equal zero is rejected at the 10 percent level ($W = 9.16$ on four degrees of freedom, $p = 0.057$). In a single cross-section, age and cohort effects cannot be separated.

C.1.3 Institutional trust

Table C.1 reports the institutional-trust index (Col. 1) and disaggregates it into its ten items (Cols. 2–11). The point estimate is negative for every institution in both panels. Controlling for schooling, six of the ten items are significant at the 5 percent level after the multiple-outcome adjustment, and two more, the local congress ($q = 0.056$) and the electoral institute ($q = 0.058$), at the 10 percent level; the exceptions are churches ($q = 0.230$) and the prosecutor's office ($q = 0.249$). The largest gaps are for the Navy (-0.47) and the Army (-0.37), on the 0–1 scale of each item.

C.1.4 Social participation

Table C.2 disaggregates participation. The estimated discontinuities in belonging to any organization (Col. 1) and in the number of memberships (Col. 2) are small and imprecise. Controlling for schooling, the point estimate for neighborhood or residents' associations is 8.0 percentage points (Col. 3), about twice the sample mean of 4.1 percent, but it is imprecisely estimated ($q = 0.200$), and no other type of membership differs (Cols. 4–7).

Table C.1: Trust in institutions across the boundary

	Institutional trust index (1)	Trust in churches (2)	Trust in local congress (3)	Trust in electoral institute (4)	Trust in political parties (5)	Trust in government (6)	Trust in judiciary (7)	Trust in police (8)	Trust in army (9)	Trust in navy (10)	Trust in prosecutor's office (11)
<i>Panel A. Without schooling control</i>											
$\mathbb{1}\{\text{Analco}\}$	-0.235*** (0.064)	-0.111 (0.116)	-0.182** (0.081)	-0.234** (0.103)	-0.178*** (0.063)	-0.259*** (0.075)	-0.195** (0.077)	-0.259*** (0.088)	-0.380*** (0.111)	-0.480*** (0.112)	-0.111 (0.085)
<i>q</i> -value	[0.004]	[0.328]	[0.049]	[0.049]	[0.018]	[0.005]	[0.035]	[0.017]	[0.005]	[0.001]	[0.221]
R ²	0.093	0.105	0.048	0.061	0.059	0.062	0.070	0.050	0.122	0.130	0.055
<i>Panel B. Controlling for years of schooling</i>											
$\mathbb{1}\{\text{Analco}\}$	-0.232*** (0.065)	-0.147 (0.114)	-0.182** (0.083)	-0.225** (0.105)	-0.178*** (0.062)	-0.258*** (0.075)	-0.187** (0.077)	-0.252*** (0.089)	-0.370*** (0.112)	-0.466*** (0.113)	-0.105 (0.087)
Years of schooling	0.002 (0.003)	-0.019*** (0.006)	0.000 (0.005)	0.005 (0.005)	0.000 (0.003)	0.000 (0.004)	0.004 (0.004)	0.004 (0.005)	0.005 (0.005)	0.008 (0.005)	0.003 (0.004)
<i>q</i> -value	[0.006]	[0.230]	[0.056]	[0.058]	[0.019]	[0.006]	[0.041]	[0.019]	[0.007]	[0.002]	[0.249]
R ²	0.094	0.139	0.048	0.064	0.059	0.062	0.073	0.053	0.125	0.136	0.057
Observations	289	289	289	289	289	289	289	289	289	289	289
Dep. var. mean	0.229	0.472	0.164	0.265	0.142	0.178	0.199	0.203	0.349	0.373	0.204
Age-Sex fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Survey weights	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: OLS estimates from the 2022 wave 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). The coefficient on the Analco indicator is the discontinuity at the boundary implied by a separate linear distance gradient on each side; unconditional means are similar across sides. All specifications include age-by-sex fixed effects and side-specific linear distance gradients, and are weighted by the survey's expansion factors. Panel A includes no schooling control; Panel B controls linearly for the respondent's years of schooling. Heteroskedasticity-robust standard errors in parentheses. Sharpened FDR *q*-values (Anderson, 2008), adjusted across all 26 social-capital outcomes within each panel, in brackets.

Table C.2: Organizational membership across the boundary

	Member of any organization (1)	Number of memberships (2)	Member of neighborhood assoc. (3)	Member of school- related org. (4)	Member of political org. (5)	Member of religious group (6)	Member of NGO/ assistance org. (7)
<i>Panel A. Without schooling control</i>							
$\mathbb{1}\{\text{Analco}\}$	-0.086 (0.100)	-0.088 (0.188)	0.066 (0.056)	-0.039 (0.050)	0.023 (0.041)	-0.076 (0.085)	-0.062 (0.063)
<i>q</i> -value	[0.329]	[0.443]	[0.268]	[0.347]	[0.399]	[0.329]	[0.328]
R ²	0.068	0.046	0.033	0.034	0.029	0.068	0.050
<i>Panel B. Controlling for years of schooling</i>							
$\mathbb{1}\{\text{Analco}\}$	-0.058 (0.101)	-0.041 (0.188)	0.080 (0.057)	-0.032 (0.050)	0.032 (0.044)	-0.064 (0.087)	-0.057 (0.063)
Years of schooling	0.015*** (0.005)	0.025*** (0.009)	0.008** (0.003)	0.004** (0.002)	0.004 (0.004)	0.007 (0.004)	0.002 (0.002)
<i>q</i> -value	[0.404]	[0.480]	[0.200]	[0.404]	[0.404]	[0.404]	[0.361]
R ²	0.098	0.073	0.054	0.046	0.037	0.079	0.054
Observations	289	289	289	289	289	289	289
Dep. var. mean	0.114	0.178	0.041	0.018	0.036	0.065	0.018
Age-Sex fixed effects	✓	✓	✓	✓	✓	✓	✓
Survey weights	✓	✓	✓	✓	✓	✓	✓

Notes: OLS estimates from the 2022 wave 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. The coefficient on the Analco indicator is the discontinuity at the boundary implied by a separate linear distance gradient on each side; unconditional means are similar across sides. All specifications include age-by-sex fixed effects and side-specific linear distance gradients, and are weighted by the survey's expansion factors. Panel A includes no schooling control; Panel B controls linearly for the respondent's years of schooling. Heteroskedasticity-robust standard errors in parentheses. Sharpened FDR *q*-values (Anderson, 2008), adjusted across all 26 social-capital outcomes within each panel, in brackets.

C.2 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 4.3.

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

	Years of schooling (1)	Trust in neighbors (2)	Neighbor vs. nonlocals trust (3)	(4)	(5)	Years of schooling (6) (7) (8) (9)			
$\mathbb{1}\{\text{Analco}\}$	-1.86 (1.13)	0.368** (0.151)	0.093 (0.064)	-0.798 (1.23)	0.172 (1.73)	-1.28 (1.25)	-2.32** (1.17)	-0.761 (1.07)	-1.81 (1.12)
Years of schooling		-0.002 (0.007)	-0.004 (0.003)						
Trust in neighbors				0.796 (0.691)					
Trust in neighbors $\times$ Analco				-1.74* (0.945)					
Neighbor vs. nonlocals trust					0.703 (1.65)				
Neighbor vs. nonlocals trust $\times$ Analco					-3.14 (2.19)				
Institutional trust index						1.67 (1.68)			
Institutional trust index $\times$ Analco						-1.89 (2.15)			
Tract-agg. neighbor trust (z)							0.775** (0.310)		
Tract-agg. neighbor trust (z) $\times$ Analco							-1.36*** (0.464)		
Tract-agg. schooling (z)								1.50*** (0.268)	
Tract-agg. schooling (z) $\times$ Analco								0.042 (0.572)	
Tract-agg. land value (log, z)									1.22*** (0.258)
Tract-agg. land value (log, z) $\times$ Analco									-0.720 (0.719)
Observations	289	289	289	289	289	289	289	289	289
Dep. var. mean	11.33	0.6549	0.5953	11.33	11.33	11.33	11.33	11.33	11.33
R ²	0.2246	0.0887	0.0812	0.2336	0.2336	0.2277	0.2476	0.3245	0.2767
Age-Sex fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimates from the 2022 wave of the local Life Quality Survey (*Jalisco Cómo Vamos*) for Guadalajara. All models include age-by-sex 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. Estimates are weighted by the survey’s expansion factors. 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 0.796, and the interaction of -1.74 implies an eastern slope of about -0.94 : among Analco residents, higher trust in neighbors goes with roughly one fewer year of schooling. The interaction has a robust p -value of 0.067 and is not statistically significant when standard errors are clustered by census tract ($p = 0.16$); it falls to -0.80 within three kilometers of the boundary. The neighbor-vs-nonlocals premium in Column (5) gives a larger difference, -3.14 against a western slope of 0.703, for an eastern slope near -2.44 . The institutional-trust index in Column (6) gives an interaction of -1.89 that is not statistically distinguishable from zero, and a western slope of 1.67, 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 4.1: the estimated discontinuity is positive for trust in neighbors (0.368) and positive but imprecise for the neighbor-versus-nonlocals premium (0.093).

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 -1.36), 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 0.04), 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 -0.72), 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 Section 5, 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.

C.3 Parishes and patron saints around the creek

Figure C.2 maps the *cuarteles* adjacent to the SJdD creek and the temples of their parishes. The patron-saint assignment used in Section 4.6 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.

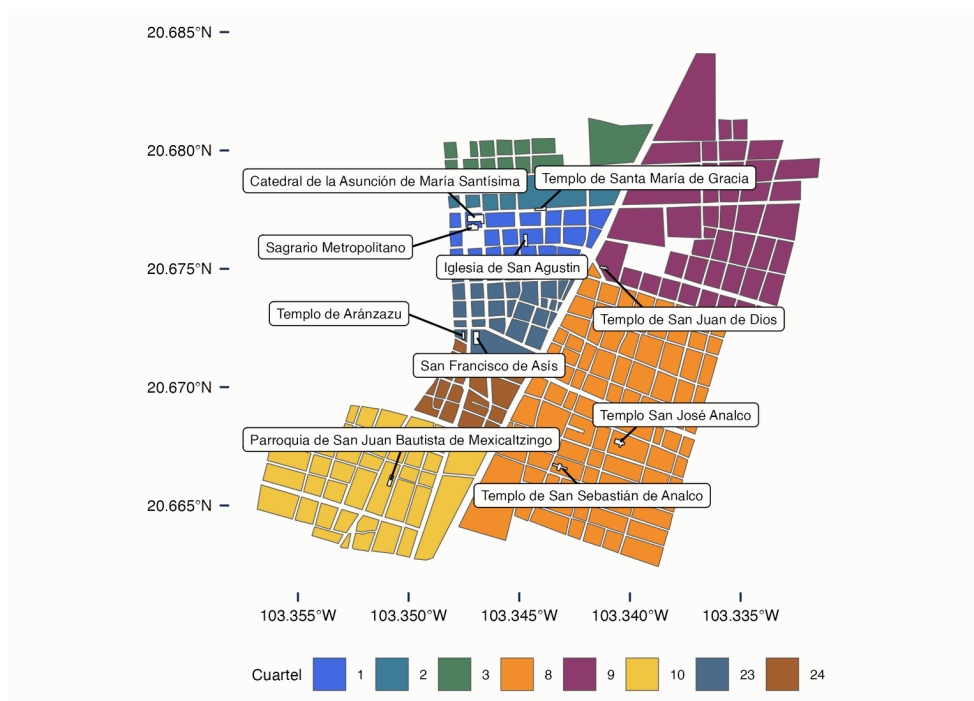
C.4 Calendar-saint naming: the santoral placebo

Figure C.3 reports the calendar-saint exercise referenced in Section 4.6. 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 4.6 is specific to the local parish.

C.5 Naming patterns: construction and additional results

This subsection documents the naming evidence summarized in Section 4.6.

Figure C.2: *Cuarteles* and parish temples around the San Juan de Dios creek

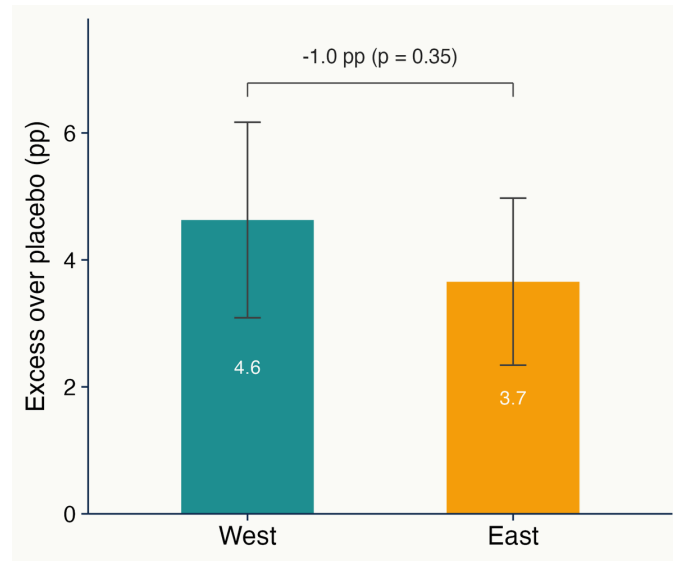


Notes: *Cuarteles* adjacent to the SJdD creek in the cartography of the Guadalajara Census Project (Anderson and Spike, 2007), colored by *cuartel* number, with the parish temples overlaid and labeled (building footprints from OpenStreetMap). Each *cuartel* is assigned the patron saints of its parish temples; the patron-saint exercise in Section 4.6 uses this assignment.

Sources. The civil *padrones* of 1821–22 and 1824 compiled by Anderson and Spike (2007) record residents’ given names, ages, and *cuarteles*. The national 10 percent sample of the 1930 census records the same fields. FamilySearch parish registers of baptisms and marriages span 1850–1949.

Patron-saint naming. 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. The Spanish west comprises the western *cuarteles* with a known parish patron, Mexicaltzingo is *cuartel* 10, under San Juan, and Analco is *cuartel* 8, under San José and San Sebastián (Figure C.2). 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. Mexicaltzingo and the Spanish west differ by –0.7 percentage points (95 percent CI [–2.6, 1.3]). 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 C.3: 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.

Checks. 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.3). 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. The gap holds in the marriage acts of 1850–1949; acts before 1850 are too sparse to support the comparison.

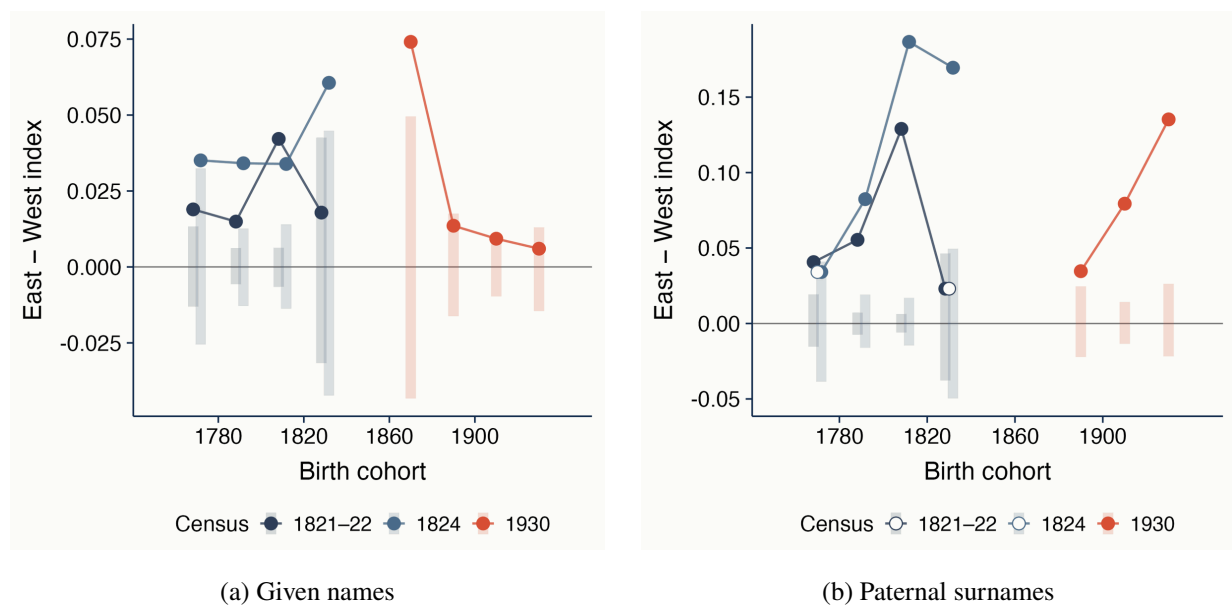
Distinct names and surnames. For each census, I compute a distinctiveness index in the spirit of [Fryer and Levitt \(2004\)](#): the probability that a resident lives east of the creek given only their first name, estimated leave-one-out and benchmarked against a permutation null. Spellings are normalized 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.

Panel (a) of Figure C.4 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 1821 census records ethnicity (*calidad*), and estimating the comparison within *calidad* cells yields the same pattern.

Panel (b) 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.

Figure C.4: Given-name and surname distinctiveness by birth cohort (1740–1939)

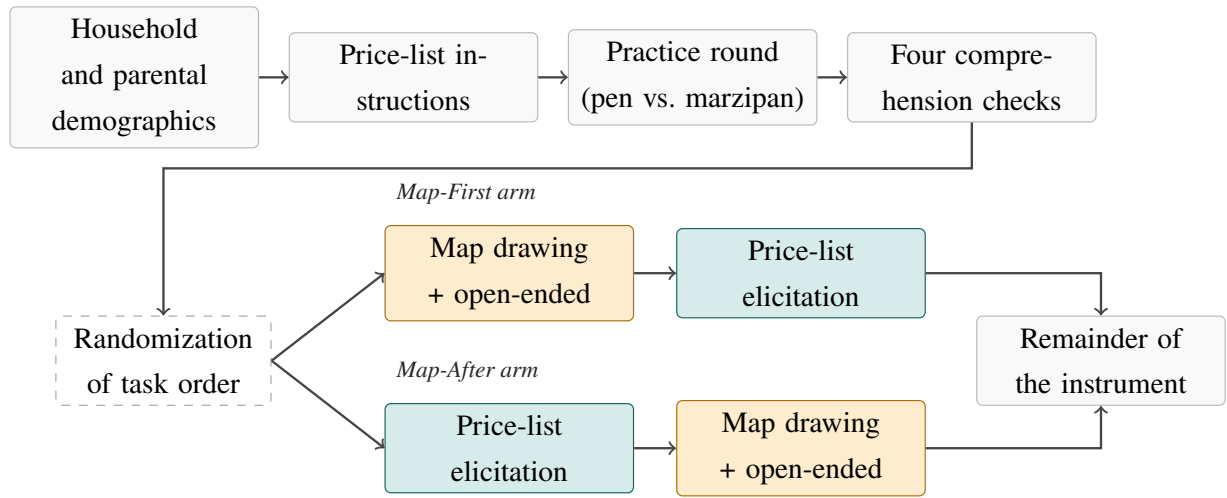


Notes: Panel (a) plots, by twenty-year birth cohort, the east–west difference in the mean leave-one-out probability that a resident lives east of the creek given their first name, estimated separately within the censuses of 1821–22, 1824, and 1930. Shaded bars are 95 percent permutation bands under random 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.

D The Incentivized Choice Experiment: Supporting Material

This appendix collects the supporting material for the incentivized choice experiment of Section 4, in the order the section draws on it. Figure D.1 sets out the order of tasks in the survey instrument. Appendix D.1 reports balance, implementation diagnostics, and influential-observation checks, and Appendix D.2 then reports the confirmatory table. Appendix D.3 characterizes what the priming activates, Appendix D.4 addresses experimenter-demand effects, and Appendix D.5 reports heterogeneity within Analco and the variation available to detect it. Appendix D.6 reports the exercise that makes the historical divide salient, and Appendix D.10 details the survey instrument and protocols.

Figure D.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.

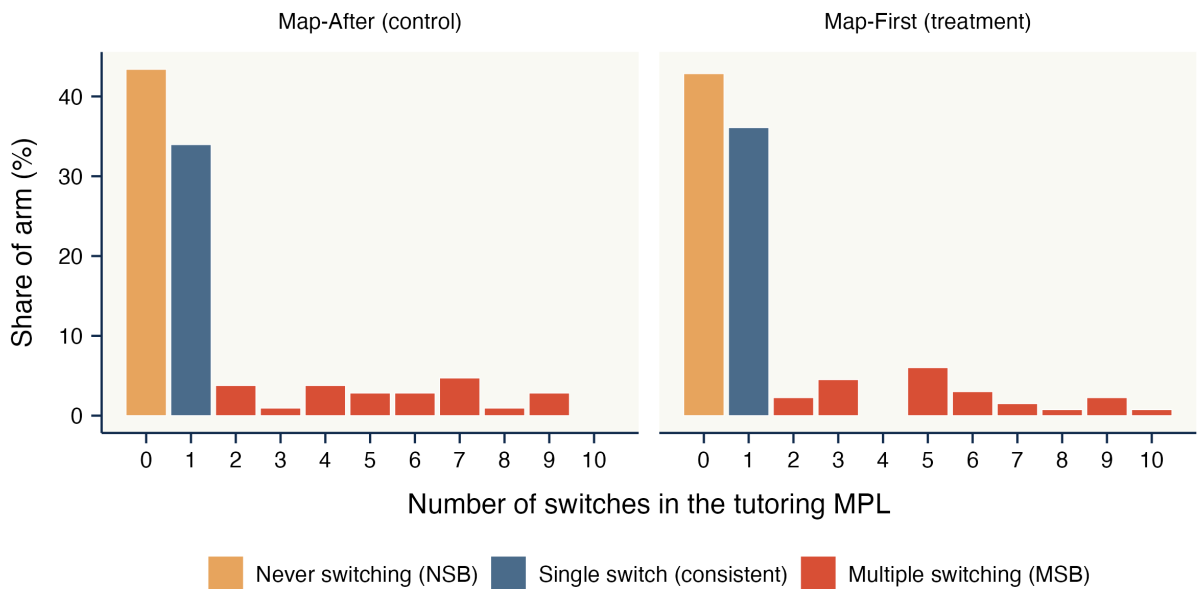
D.1 Balance, diagnostics, and influential observations

Covariate balance. Table D.1 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 absolute distance to it, parent and grandparent education, household composition, income bracket, and price-list comprehension — and none differs significantly across arms after a sharpened false-discovery-rate correction (Anderson, 2008). Individually, only the absolute distance to the boundary differs at the 5 percent level, by 209 meters ($p = 0.049$), and the specifications of Table D.4 already control for that distance with a separate slope on each side. 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 D.2 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 D.2 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.

Figure D.2: 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.

Randomization inference and influential observations. Tables D.3 and D.4 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 D.4, 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 $\times$ 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 D.3 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 $[-78.0, -62.3]$ pesos for willingness to pay, so no single household drives the result.

Interviews before registration, distance, and the west side. The registration is dated May 14, 2026, and 23 households, 17 on the east side and 6 on the west, were interviewed before that date. Excluding them gives an east-side effect of -0.124 on the tutoring share (randomization-inference $p = 0.014$) and an interaction of -0.297 ($p = 0.039$), against -0.123 and -0.267 in the full sample. The specification of Table D.4 controls for absolute distance to the boundary with a separate slope on each side, because the signed distance interacted with the Analco indicator spans it exactly. Replacing the linear terms with side-specific quadratics or with side-by-500-meter-bin fixed effects leaves the east-side effect between -0.119 and -0.127 . Restricting the sample to households within two kilometers of the boundary, 126 of them on the east side, gives -0.101 ($p = 0.14$). On the west side, Map-First households have lived in their *colonia* 10.6 fewer years than Map-After households (randomization-inference $p = 0.040$) and report belonging 1.4 points lower on the seven-point scale ($p = 0.014$), while no local-ties measure differs across arms on the east side. The west-side estimate is 0.144 in the full sample, 0.143 without the respondents who inferred the purpose of the study, and 0.099 when that inference is controlled for, and it is never distinguishable from zero. The randomization-inference p -value of the difference between sides ranges from 0.039 to 0.11 across these variants.

D.2 Confirmatory results

Table D.4 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 D.4.

Inference. Map-First was randomized within side of the historical boundary, so the permutation distribution holds fixed the number treated on each side. The reported randomization-inference

p -values are the share of 50,000 permutations producing an estimate at least as large in magnitude as the observed one; Table D.4 also reports conventional robust standard errors.

Magnitude. Using Glass’s Δ , the east-side effects correspond to 0.33 standard deviations for the tutoring share and 0.30 for willingness to pay. The MXN 75 decline in willingness to pay is about one quarter of Mexico’s 2026 general daily minimum wage, set by the *Comisión Nacional de los Salarios Mínimos* at MXN 315.04 from January 1, 2026 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.

Construction of the main-text figure. In Figure 5, the Map-After bar is the unadjusted control mean on each side. The Map-First bar adds to it the estimated Map-First effect on that side from the interacted specification in columns (2) and (5) of Table D.4: the Map-First coefficient on the west side, and the sum of the Map-First coefficient and its interaction with Analco on the east side. The whiskers are the randomization-inference interval for that effect, shifted by the control mean, which is treated as fixed. The bracket between the two Map-After bars compares the unadjusted control means across sides. The side of the boundary is not randomized, so its p -value does not come from randomization inference: it uses robust standard errors for the tutoring share and the respondent-level bootstrap of both steps for willingness to pay. On the east side, the implied Map-First levels are 45.0 percent and MXN 102, the values reported in the main text; the specification estimated on Analco alone, in columns (3) and (6), gives effects of -11.7 percentage points and MXN -71 .

Latent willingness to pay. The pre-registered willingness-to-pay measure (Table D.4) is the two-step latent estimator of Jack et al. (2022), registered as a primary outcome in AEARCTR-0018600. Table D.5 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 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.

D.3 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.

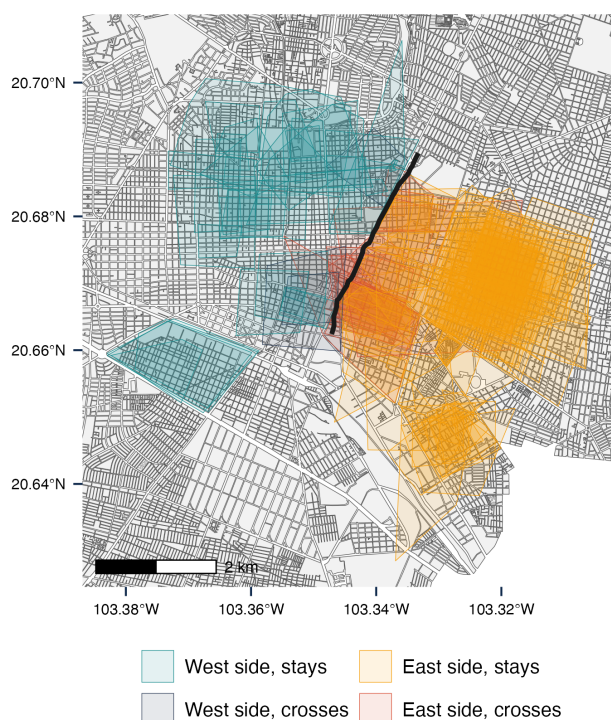
The geometry of the drawn barrio. Table D.6 summarizes the polygons. Each traced barrio is represented by the convex hull of its vertices. The neighborhoods people draw are small (a median of just over half a square kilometer — a few blocks), and home-centered: for 88% of respondents the geolocated home falls *inside* their own drawing. Only 11% of the drawn barrios cross the historical boundary, the SJdD creek, today Calzada Independencia (Figure D.3). 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.

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 D.10.1), 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 D.7 reports the shares by side of the boundary.

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 4.3 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

Figure D.3: 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-side respondents whose drawing stays on their side and navy for those whose drawing crosses, amber for east-side respondents whose drawing stays and brick for those whose drawing crosses. Polygons with obvious drawing errors, such as a misplaced distant vertex, are omitted.

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.

D.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 D.10.1) — 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 D.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 side rather than in Analco. The Map-First $\times$ 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 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 sides, since the guessing moves on the side where behavior does not.

D.5 Heterogeneity within Analco, and the variation available

If the effect ran through household resources or inherited human capital rather than through the neighborhood, it should be concentrated among poorer or less-educated households. Table D.9 adds each moderator and its interaction with Map-First to the Analco-subsample specification. No interaction is significant. The effect is present among households with above-median income (-0.176), among those where a parent completed high school (-0.129), and among those that are both (-0.207), and in each case it is at least as large as among the households without those advantages.

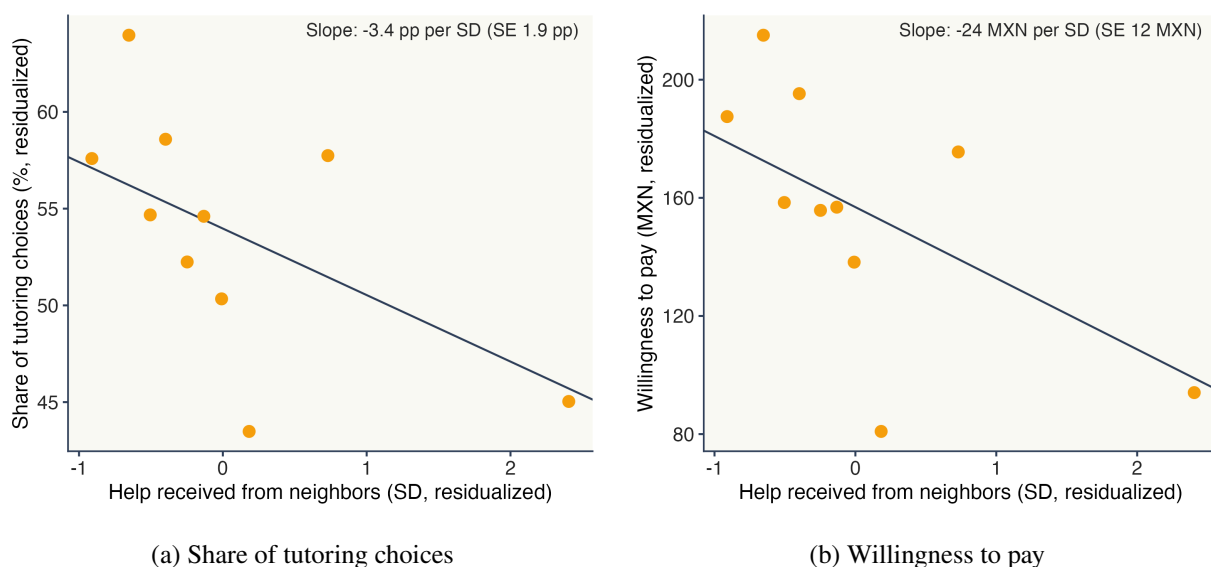
A null interaction has two possible sources: the effect may not vary with the moderator, or the moderator may not vary in the sample. The first column of Table D.9 reports the mean and dispersion of each moderator among Analco households to separate them. Income, parental schooling, and grandparental schooling are dispersed on the eastern side, and about as dispersed as on the western side. The local-ties index varies more in Analco than in the west, with a standard deviation of 1.09 against 0.26, so the eastern sample is not a homogeneous block of households whose social relations are concentrated locally.

What the design cannot do is detect moderate interactions. With 193 Analco households, the interaction terms carry standard errors around 0.11 for a median split, so the smallest interaction detectable with 80 percent power is about 0.30, roughly two and a half times the effect being split. These rows therefore rule out the case where the effect exists only among constrained households,

and they do not establish that the effect is uniform.

The individual measures of local exchange behave like the index. Interacting Map-First with help received, hours with neighbors, relatives in the *colonia*, years of residence, the relative *barrio* premium, and belonging gives a joint randomization-inference p -value of 0.91, and interacting it with the average of each measure among the other sampled households of the same *colonia* gives 0.25. Outside the interaction, help received from neighbors is associated with a tutoring share 3.4 percentage points lower per standard deviation in the full east-side sample ($p = 0.074$; Figure D.4) and 5.4 points lower in the Map-After arm ($p = 0.19$); no other measure predicts demand. The measures barely cluster by *colonia*, with intraclass correlations between -0.06 and 0.02 for all but the share of time spent with neighbors, so the sample cannot separate a payoff attached to the place from one attached to the household.

Figure D.4: Help received from neighbors and demand for tutoring on the east side



Notes: Binned scatter plots for the 193 households on the east side, in ten bins of equal size. The horizontal axis is the number of times the household received help from neighbors in the previous year, in standard deviations. The vertical axis is the share of the eleven choices in which tutoring is selected in Panel (a) and latent willingness to pay in MXN in Panel (b). Both variables are residualized on the Map-First assignment and on the controls and fixed effects of the main specification (Table D.4), and their sample means are added back. The line is the linear fit, and its slope is the association reported in the text. This analysis was not pre-registered. Help received is measured after the experimental block; the treatment does not change it ($p = 0.76$).

D.6 Making the historical divide salient: 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 D.10 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 side 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 sides separately and no cross-side test. A displacement in Analco alongside a null on the west side does not establish that the two sides differ, and with 46 west-side respondents the design has little power to tell them apart.

Nearly 80% of respondents report hearing or using the expression. Among Analco respondents, references to money or the need to work fall 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. Controls for the earlier map assignment, interview length, and respondent demographics leave the estimates similar. The exercise identifies an effect of making the historical divide salient on explanations. It does not identify the historical assignment's effect on those explanations or any component of the local cost in Section 4.2, and it is not a replication of the tutoring experiment, whose treatment and outcome differ.

D.7 Survey modules: additional detail

This subsection collects detail on the survey comparisons summarized in Section 4.5 and reported in Table 3.

Types of help from neighbors. The help index in Panel B of Table 3 sums three items asking how often, over the previous twelve months, the respondent received each type of help from neighbors, with frequency bins converted to counts. The most common type combines childcare, support during an illness, and household emergencies: 31% of Analco respondents and 20% of

west-side respondents report receiving it at least once, and the difference in the number of times is imprecise. Loans of money or goods are reported by 13% and 11%. Help finding work shows the sharpest difference, reported by 17% of Analco respondents against 2% on the west side. These are unadjusted comparisons, and the first item does not separate childcare from the other two forms of help it groups.

Returns and the reservation salary. Parents on both sides expect returns well above the realized 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.20). The reservation-salary 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, and respondents are asked how much the young person would need to earn per month at age 30 for university to have been worthwhile. The MXN 5,300 difference is roughly 45% of the university wage premium respondents expect.

Open-ended descriptions of the neighborhood. Most answers on both sides refer to streets, landmarks, and everyday routes. A minority define the neighborhood by the people in it, and this minority is larger in Analco, 15% against 7% on the west side (28 of 193 and 3 of 46 answers), although the difference is imprecise. The original answers quoted in the main text are “Dónde me conocen y conozco a los vecinos,” “Hasta donde la gente se conoce del barrio,” and “Dónde hay más gente de confianza.” Other Analco respondents draw the neighborhood along “the streets that count for organizing the patron-saint festivities” (“Son las calles que se toman en cuenta para la organización de las fiestas patronales”). The answers were given after drawing in both treatment arms, and these examples were selected to illustrate the category.

Allocation exercises. The allocations follow Enke et al. (2023) and Cappelen et al. (2025). In the first, respondents 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 respondent keeps nothing, so each allocation reflects the relative weight placed on the two recipients. A third question asks how much of an unexpected MXN 500 windfall the respondent would donate to a good cause. The relative *barrio* premium in Panel A of Table 3, which scales the first allocation by the second, is driven by a few respondents who give little to the Guadalajara stranger, and symmetric versions of the same comparison are small and imprecise.

Aspirations, norms, and attachment. Aspirations and expectations refer to a randomly selected focal child and control for the child’s sex, age, and current school attendance (Panel C).

Following [Bursztyn and Yang \(2022\)](#), respondents report their own and their perceived neighbors' agreement with the statement that young people should work rather than continue studying; neither differs across the boundary, nor does pluralistic ignorance (Panel E). Stated attachment covers feelings of belonging, preferences over living elsewhere, and effort to have children know the neighbors (Panel F).

D.8 Magnitude checks

Two comparisons relate the survey answers to the experiment and to the census gap. They compare magnitudes across exercises that use different goods, horizons, and contrasts, and they are not structural estimates.

The reservation salary and willingness to pay. East-side parents report a reservation salary of MXN 36,363 per month at age 30, against MXN 31,043 on the west side, a difference of MXN 5,319 (Table [3](#)). Respondents expect a university graduate to earn MXN 11,704 more per month than a high-school graduate, so the difference is 45 percent of the premium they expect. In the experiment, Map-First lowers willingness to pay for tutoring on the east side from MXN 176 to MXN 102, or by 42 percent (Table [D.4](#)).

The reservation salary and years of schooling. Among east-side parents, 18.1 percent expect a university graduate to earn at least their reservation salary. If every east-side parent's reservation salary were MXN 5,319 lower, the share would be 36.8 percent, 18.7 percentage points higher. If the children of these marginal parents completed four more years of schooling, average schooling on the east side would be 0.75 years higher (0.187×4), against a boundary gap of 0.8 years in the 2020 Census (Table [B.1](#)). The calculation assumes that each marginal child would complete a four-year degree, and it does not apportion the census gap among its sources.

D.9 Household characteristics of the experimental sample

Table [D.11](#) compares the households on the two sides of the boundary. Household income, household size, parental and grandparental schooling, the share of close relatives living in the *colonia*, years of residence, and whether a parent lived there do not differ after the correction for multiple comparisons. Among respondents who work, a written contract is, if anything, more common in Analco, and self-employment is similar on the two sides; these two rows also control for the respondent's schooling.

D.10 Incentivized choice experiment: additional design detail

This subsection collects further detail on the incentivized choice experiment beyond Section [4.3](#).

D.10.1 Survey instrument

The interview has a median duration of 26.5 minutes, and the instrument 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 D.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.

D.10.2 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.

D.10.3 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: an incorrect answer triggers the relevant instructions before the question is re-administered, and the original answer remains recorded. In the main sample the mean score is 3.8 out of 4. The binding row is determined by a numbered slip that the respondent draws from a sealed envelope; the binding-row procedure and the one-in-ten payment probability are explained before the task, and all households retain the MXN 100 participation payment regardless of the experimental outcome. Pilot results show 92% comprehension on the first pass and 100% after one round of

clarification.

Table D.1: Covariate balance across the Map-First treatment

	Map-After (control)	Map-First (treatment)	Difference (T – C)	<i>p</i> -value	<i>q</i> -value	<i>N</i>
<i>Panel A. Measured before the randomized block</i>						
$\mathbb{1}\{\text{Analco}_i\}$	0.840	0.782	-0.058	—	—	239
Absolute distance to boundary (m)	1465	1674	209	0.049	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
<i>Panel B. Measured after the randomized block</i>						
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 D.10.1, 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 whether the treatment shifted that choice. Columns (1)–(2) report arm means and column (3) their difference (treatment minus control); column (4) is the randomization-inference *p*-value for that difference and column (5) the corresponding Benjamini–Hochberg FDR *q*-value across all tests in the table. Every *p*-value comes from permuting the Map-First label within side of the historical boundary — the randomization stratum of the experiment — holding fixed the number of treated households on each bank, and recomputing the difference. The side indicator itself is the stratum, so it is identical under every admissible permutation and admits no test; its row is left blank. The distance row is in absolute value; the specifications of Table D.4 control for that distance with a separate slope on each side of the boundary. No individual difference in either panel survives the FDR correction. *N* varies across rows because some parental-education and income items are missing for a minority of respondents.

Table D.2: Incentivized choice experiment: implementation diagnostics

	Map-After (control)	Map-First (treatment)	All	Difference <i>p</i> -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.

Table D.3: Robustness of the incentivized-choice-experiment estimates: leave-one-out

	Estimate	Leave-one-out range
<i>Panel A. Share of tutoring choices</i>		
Pooled effect (H1)	-0.077	[-0.090, -0.067]
Map-First $\times$ Analco (H2)	-0.267	[-0.297, -0.195]
Priming effect in Analco	-0.117	[-0.130, -0.103]
<i>Panel B. Willingness to pay (MXN pesos)</i>		
Pooled effect (H1)	-44.7	[-53.1, -38.5]
Map-First $\times$ 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 D.4: the pooled Map-First effect (H1), the Map-First $\times$ Analco interaction (H2), and the priming effect in Analco, estimated on the east-side subsample. “Estimate” reproduces the point estimate; its inference, conventional and by randomization, is reported in Table D.4. 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 D.4. The three rows replicate columns (1)–(3) of that table: the pooled effect, the Map-First $\times$ Analco interaction, and the east-side 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.

Table D.4: 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)
$\mathbb{1}\{Map-First_i\}$	-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		—			—	
$\mathbb{1}\{Map-First_i\} \times \mathbb{1}\{Analco_i\}$		-0.267**			-172.2*	
Robust		(0.137)			(102.9)	
Randomization inference		(0.129)			(88.4)	
Sample	Full	Full	East side	Full	Full	East side
Enumerator FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
LASSO controls	✓	✓	✓	✓	✓	✓
Distance to boundary	✓	✓	✓	✓	✓	✓
<i>p</i> -value: east-side effect = west-side 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.

Table D.5: 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 D.4; 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-side subsample estimate reported in Table D.4 and Table D.3. Both measures deliver the same conclusion: a negative priming effect concentrated in Analco.

Table D.6: Geometry of the neighborhood respondents drew

	All	Analco	West side
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.

Table D.7: What respondents had in mind when drawing their neighborhood

	All	Analco	West side
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.

Table D.8: Robustness to experimenter-demand effects

<i>Panel A. Robustness of the priming effect in Analco</i>			
	(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
<i>Panel B. Effect of the treatment on the inferred topic (linear probability models)</i>			
	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 (11m-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 D.4, 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 D.4 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 side 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 D.9: Heterogeneity of the Map-First effect within Analco, and the variation available

Moderator	Analco mean (SD)	Map-First effect		Interaction			<i>N</i>
		Low / at 0	High	Estimate	RI <i>p</i>	MDE	
Local-ties index (PCA), per SD	0.08 (1.09)	−0.117 (0.054)	—	−0.004 (0.045)	0.942	0.126	193
Local-ties index above median	0.54	−0.116 (0.087)	−0.124 (0.067)	−0.007 (0.111)	0.953	0.310	193
Income bracket (1–10), per bracket	4.27 (2.29)	−0.046 (0.122)	—	−0.014 (0.026)	0.573	0.073	174
Income bracket above median	0.41	−0.072 (0.074)	−0.176 (0.090)	−0.104 (0.114)	0.347	0.320	174
Parental years of schooling, per year	12.40 (3.45)	−0.174 (0.205)	—	0.005 (0.015)	0.774	0.043	193
A parent completed high school	0.68	−0.093 (0.096)	−0.129 (0.059)	−0.035 (0.107)	0.752	0.300	193
A grandparent completed high school	0.37	−0.123 (0.065)	−0.115 (0.089)	0.008 (0.107)	0.931	0.300	190
Above-median income and a parent with high school	0.37	−0.059 (0.071)	−0.207 (0.093)	−0.148 (0.116)	0.193	0.324	174

Notes: Each row adds a moderator *Z* and its interaction with Map-First to the Analco-subsample specification of Table D.4: the share of tutoring choices on Map-First, controls selected by LASSO, and enumerator and day fixed effects, with HC1 standard errors. For binary moderators the two effect columns are the Map-First effect at $Z = 0$ and at $Z = 1$; for continuous moderators the first column is the effect at the sample mean and the second is omitted. The first column reports the mean of the moderator among Analco households, with its standard deviation in parentheses for continuous moderators, and is there to show how much variation the interaction can use. MDE is the interaction size the design would detect with 80% power at the 5% level, 2.8 times the standard error. Randomization-inference *p*-values redraw Map-First within the Analco subsample, holding the moderator and the controls fixed, with 2,000 permutations. The local-ties index is the first principal component of the survey-module measures that differ across the boundary: allocation to the *barrio* stranger, giving strictly more to that stranger, the relative *barrio* premium, times helped by neighbors, hours with neighbors, and the reservation salary for university. Higher values indicate social relations more concentrated locally. It is measured after the randomized block, so its interaction is descriptive. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table D.10: Boundary salience and explanations of educational non-continuation

	Share mentioning		Prime effect		Adjusted	
	Not primed	Boundary primed	Effect	<i>q</i>	Effect	<i>q</i>
<i>Panel A. Analco, historically excluded bank (N = 193)</i>						
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
<i>Panel B. West side (N = 46)</i>						
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. This analysis is exploratory. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table D.11: Household characteristics of the experimental sample

	Mean west side	Mean Analco	Analco coefficient	q
<i>Panel G. Household income, family schooling, and roots in the colonia</i>				
Monthly household income [†]	17,400	15,705	−0.100 (0.140)	1.00
Monthly household income per member [†]	4,149	3,824	−0.065 (0.148)	1.00
Household size	4.48	4.67	0.10 (0.24)	1.00
Years of schooling, more-schooled parent	13.15	12.40	−0.69 (0.52)	1.00
A parent completed high school	0.739	0.679	−0.050 (0.071)	1.00
Years of schooling, respondent's father	9.58	8.08	−1.83** (0.90)	1.00
Years of schooling, respondent's mother	7.77	8.08	0.16 (0.71)	1.00
Schooling gain over the more-schooled grandparent (years)	3.52	3.27	0.04 (0.77)	1.00
Share of close relatives living in the colonia	0.141	0.196	0.050 (0.036)	1.00
Years lived in the colonia	25.70	25.17	−0.74 (2.74)	1.00
A parent lived in the colonia	0.478	0.580	0.091 (0.083)	1.00
Works with a written contract (respondents who work)	0.294	0.433	0.142 (0.098)	1.00
Self-employed (respondents who work)	0.412	0.370	−0.060 (0.101)	1.00

Notes: The first two columns report unadjusted means for west-side and Analco respondents. The third column reports the coefficient on the Analco indicator, controlling for Map-First assignment and respondent sex, age, and age squared. The household rows omit the respondent-education fixed effects, which would absorb parental schooling mechanically; the two informality rows, estimated among respondents who work, include them. Income is the midpoint of the reported bracket; years of schooling are assigned to completed levels; the more-schooled parent is the respondent or the partner. The share of close relatives (siblings, uncles and aunts, cousins) living in the colonia maps the five answer categories to 0, 0.25, 0.5, 0.75, and 1; the colonia is the area the respondent drew on the map. A written contract and self-employment come from the question “Do you have a formal work contract?”, whose answers distinguish a written contract, work without a written contract, and self-employment. HC1 robust standard errors are in parentheses. The sample is smaller where income, a grandparent’s schooling, or employment is not reported. Outcomes marked [†] are estimated by Poisson pseudo-maximum-likelihood; all other outcomes are estimated by OLS in levels. The q column reports sharpened false-discovery-rate values (Anderson, 2008) computed across the rows of the table. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

E External Validity: Supporting Material

This appendix collects the supporting material for Section 4.7, 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 distance; the complementary distant-ties share is the fraction reaching far. Both measure how geographically local a place’s friendship ties are, although only for ties recorded on Facebook. Table E.1 covers the world’s subnational regions (schooling from the Global Data Lab), Table E.2 the NUTS2 regions of Europe (tertiary attainment from Eurostat, the developed-economy counterpart absent from the Global Data Lab sample), and Table E.3 the municipalities of Mexico (2020 Census), with the scope condition estimated against, respectively, ethnic segregation and block-level SES segregation. Table E.4 reports U.S. ZIP-code areas (ACS 5-year), where county residential segregation moderates the gradient. Appendix E.1 details the World Values Survey sample behind the individual-level evidence.

Table E.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*** (0.2357) [0.3394]	-2.806*** (0.3497) [0.4515]	-1.706*** (0.3540) [0.4381]	-3.130*** (0.4873) [0.5399]	-2.442*** (0.2348) [0.3471]	-2.583*** (0.3481) [0.4369]	-1.473*** (0.3686) [0.4292]	-2.752*** (0.5115) [0.5657]	-1.776*** (0.2577) [0.3685]	-1.880*** (0.4049) [0.4597]	-1.173*** (0.4396) [0.4641]	-2.292*** (0.5432) [0.5600]
Distant social ties (share ≥ 500 km)		-0.3266 (0.5701) [0.9035]	0.7916 (0.5105) [0.7712]	-0.2984 (0.6426) [0.8044]		-0.2947 (0.5616) [0.7963]	0.7895 (0.5320) [0.7166]	-0.1324 (0.6867) [0.7237]		-0.2171 (0.5920) [0.8018]	0.3383 (0.5764) [0.7645]	-0.4706 (0.6643) [0.7150]
Localization of social ties $\times$ Ethnic segregation				-0.4000* (0.2398) [0.2639]				-0.3980 (0.2757) [0.3104]				0.0820 (0.2842) [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 4.7 rests on Table E.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.

E.1 The World Values Survey sample

Table E.5 reports the countries available for the individual-level analysis of Section 4.7. 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 E.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*** (2.392)	-30.85*** (4.264)	-50.92*** (6.315)
Distant social ties (international share)		-14.61 (9.792)	-12.28 (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 E.1 and Table E.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.

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 E.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.

Specification. The specifications include region-by-year, region-by-settlement-size, sex-by-age, country-of-birth, religion, ethnic-group, and parental-education fixed effects, so they control flexibly for mother’s and father’s education. They exclude respondent occupation, which is realized after schooling. For educational mobility, forty of the forty-two countries have the outcome, and the same seven countries leave the main estimates. 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 larger out-group coefficient is

Table E.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*** (0.1508)	-3.373*** (0.1875)	-1.487*** (0.1977)	-1.421*** (0.2038)	-0.1003*** (0.0130)	-0.1282*** (0.0163)	-0.1255*** (0.0222)	-0.0949*** (0.0221)
Distant social ties (international share)		1.807* (0.9858)	-2.704*** (0.8620)	-3.341*** (0.8395)		-0.2979*** (0.0800)	-0.2066** (0.0873)	-0.2528*** (0.0875)
SES segregation (block)				0.3760*** (0.0602)				0.0261*** (0.0055)
Localization of social ties × SES segregation (block)				-0.6209*** (0.1221)				1.57 × 10 ⁻⁵ (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 E.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.

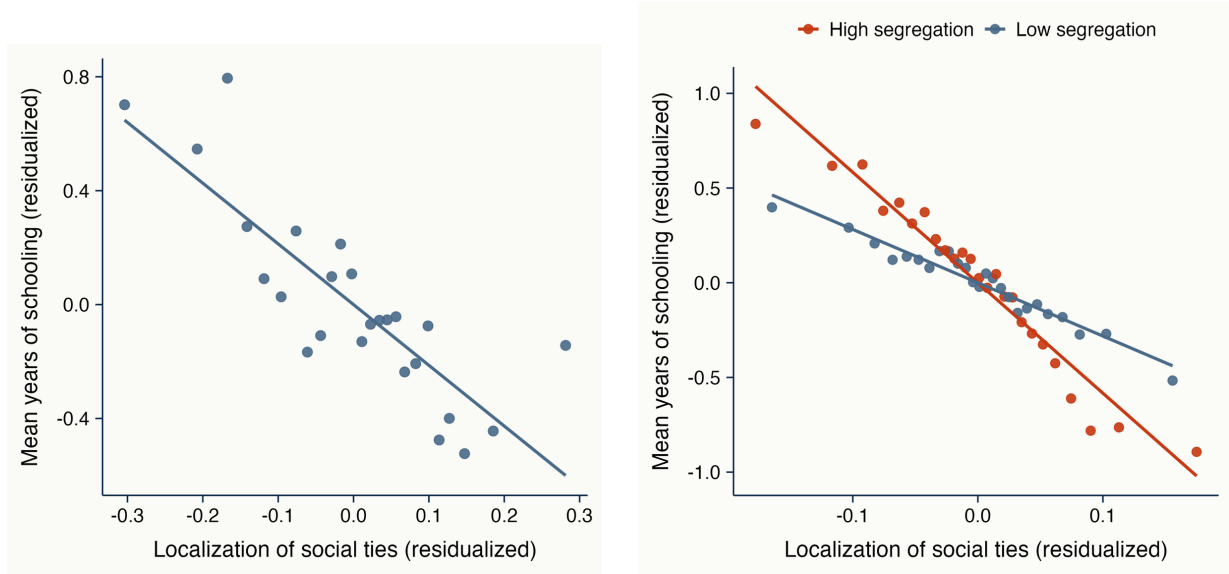
Table E.4: Localization of social ties and schooling across U.S. ZIP-code areas (ZCTA)

	Mean yrs. schooling				Share BA+			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Localization of social ties (share ≤25 km)	-3.539*** (0.0819) [0.1864]	-1.751*** (0.1104) [0.2202]	-2.500*** (0.1254) [0.2360]	-2.376*** (0.1234) [0.2153]	-0.5356*** (0.0115) [0.0271]	-0.2535*** (0.0159) [0.0333]	-0.3436*** (0.0184) [0.0369]	-0.3185*** (0.0179) [0.0322]
Distant social ties (share ≥500 km)		4.055*** (0.2122) [0.4248]	4.306*** (0.2428) [0.4582]	4.447*** (0.2431) [0.4493]		0.6400*** (0.0325) [0.0627]	0.6259*** (0.0370) [0.0665]	0.6530*** (0.0369) [0.0646]
Localization of social ties × Racial segregation				-0.7887*** (0.0758) [0.1490]				-0.1653*** (0.0114) [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.

related to Chetty et al. (2022), who find that upward mobility is associated with connectedness to higher-SES contacts.

Figure E.1: 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 E.1 and E.4.

Table E.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 E.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 E.6: Neighborhood trust premium by segregation regime: sensitivity to the fixed-effect set

	Low ethnic segregation	High ethnic segregation	Observations	Countries
<i>Panel A. Years of education</i>				
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
<i>Panel B. Upward educational mobility</i>				
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
<i>Panel C. Dropout before upper secondary</i>				
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.

F The Framework: Derivations and Proofs

This appendix derives the two models of Section 5 step by step, using the notation introduced there. Appendix F.1 sets out the timing, Appendix F.2 solves the static model, and Appendix F.3 solves the dynamic model.

F.1 Environment and Timing

Time is indexed by generations $t = 0, 1, 2, \dots$. In each generation, every neighborhood $\ell \in \{E, W\}$ has a continuum of households of mass one, each formed by an adult and a child aged n . Some adults are disposed to participate in local exchange, and $q_{\ell t}$ denotes the share of adults who participate. Each generation unfolds in five steps.

1. Adults observe the participating share $q_{\ell t}$, the quality of outside channels $Z_{\ell t}$, and educational provision $P_{\ell t}$.
2. Each household disposed to participate decides whether it can afford to do so, which determines θ_i .
3. Each household chooses whether to invest in its child's education, e_i .
4. Each adult chooses how much effort τ to put into transmitting its disposition to the child.
5. Children become the adults of generation $t + 1$, with the disposition they acquired in step 4 and parental human capital $h_{i,t+1} = e_{i,t}$.

Steps 2 and 3 form the static model, and steps 4 and 5 form the dynamic model. Within a generation, every household takes $q_{\ell t}$ as given.

F.2 The Static Model

Step 1: The value of local exchange. Each household needs one unit of services per generation, such as care, small loans, and help in emergencies. Obtained from a neighbor through reciprocal exchange, a unit is worth $R > 0$. Obtained through outside channels, it is worth $Z_{\ell}(1 - q_{\ell})$. The factor $Z_{\ell} \geq 0$ is the quality of outside channels, and the factor $1 - q_{\ell}$ captures their thickness: as in Kranton (1996), outside channels are worse when fewer households use them. The surplus of local exchange over outside channels is therefore

$$\Delta(q_{\ell}, Z_{\ell}) = R - Z_{\ell}(1 - q_{\ell}), \quad \frac{\partial \Delta}{\partial q_{\ell}} = Z_{\ell} \geq 0, \quad \frac{\partial \Delta}{\partial Z_{\ell}} = -(1 - q_{\ell}) \leq 0.$$

A household that draws on local exchange with intensity β_i obtains $\beta_i \Delta(q_{\ell}, Z_{\ell})$ from participating. The static results use only the signs of the two derivatives; the linear form matters only for Proposition 6.

Step 2: Participation. A household disposed to participate does so if the value of participating covers its cost, $\beta_i \Delta(q_\ell, Z_\ell) \geq c(y_i)$, with $c'(y) < 0$. Participation is decided before education, as a standing arrangement with neighbors, so the comparison uses the value of participation before any investment in the child. Households not disposed to participate, and disposed households that cannot afford to, have $\theta_i = 0$. Because c decreases with resources, the poorest households are the most likely to have $\theta_i = 0$.

Step 3: The investment decision. Given θ_i , the household's payoff is (4):

$$U_i(e_i) = (\rho - C_i) e_i + \theta_i [(1 - \pi_n e_i) \beta_i \Delta_\ell - c(y_i)], \quad \Delta_\ell \equiv \Delta(q_\ell, Z_\ell).$$

Evaluating it at the two choices gives

$$U_i(1) = \rho - C_i + \theta_i [(1 - \pi_n) \beta_i \Delta_\ell - c(y_i)], \quad U_i(0) = \theta_i [\beta_i \Delta_\ell - c(y_i)].$$

The difference is

$$U_i(1) - U_i(0) = \rho - C_i - \theta_i \pi_n \beta_i \Delta_\ell = x_i - \Pi_i,$$

where $x_i \equiv \rho - C_i$ and $\Pi_i \equiv \theta_i \pi_n \beta_i \Delta_\ell$, as in (5). The participation cost $c(y_i)$ appears in both $U_i(1)$ and $U_i(0)$ and cancels.

Step 4: Salience. When choosing, parents weigh the local-exchange term in brackets by $a(s) \in (0, 1]$ instead of one, with $a'(s) > 0$. The decision payoff is therefore

$$\tilde{U}_i(e_i) = (\rho - C_i) e_i + a(s) \theta_i [(1 - \pi_n e_i) \beta_i \Delta_\ell - c(y_i)],$$

and repeating step 3 gives $\tilde{U}_i(1) - \tilde{U}_i(0) = x_i - a(s) \Pi_i$. The household invests if $x_i \geq a(s) \Pi_i$, which is (2). Let $s_0 < s_1$ denote salience in the Map-After and Map-First arms, and $a_0 \equiv a(s_0) < a_1 \equiv a(s_1)$.

Step 5: Demand and willingness to pay. Let $F_\ell(\cdot \mid h, y)$ be the distribution function of x_i among households in neighborhood ℓ with $(h_i, y_i) = (h, y)$, with density $f_\ell(\cdot \mid h, y)$, and assume that the idiosyncratic cost ε_i is independent of (θ_i, β_i) given (h_i, y_i) . Educational demand is the share of households that invest,

$$D_\ell(s) = \Pr(x_i \geq a(s) \Pi_i) = 1 - \mathbb{E}[F_\ell(a(s) \Pi_i \mid h_i, y_i)].$$

Tutoring is a marginal investment of size $\delta > 0$ whose return, cost, and local cost are proportional to δ . A household is indifferent between the tutoring session and a sum of money W when $\delta [x_i - a(s) \Pi_i] = \kappa W$, where $\kappa > 0$ is the marginal utility of money. Willingness to pay is therefore

$$W_i(s) = \omega [x_i - a(s) \Pi_i], \quad \omega \equiv \delta / \kappa.$$

Results.

Proposition 1 (Saliency). *Demand falls with salience, $D_\ell(s_1) - D_\ell(s_0) = -\Pr(a_0\Pi_i \leq x_i < a_1\Pi_i) \leq 0$, and does not change if $\Pi_i = 0$ for all households in ℓ . Average willingness to pay falls by $\mathbb{E}[W_i(s_0) - W_i(s_1)] = \omega(a_1 - a_0) \mathbb{E}[\Pi_i]$, which increases with q_ℓ and decreases with Z_ℓ .*

Proof. A household invests in the Map-First arm only if $x_i \geq a_1\Pi_i$, which implies $x_i \geq a_0\Pi_i$, because $a_1\Pi_i \geq a_0\Pi_i$. The households that invest in the Map-After arm but not in the Map-First arm are those with $a_0\Pi_i \leq x_i < a_1\Pi_i$, which gives the first expression. If $\Pi_i = 0$, the interval is empty. For willingness to pay, $W_i(s_0) - W_i(s_1) = \omega(a_1 - a_0)\Pi_i$; taking expectations gives the second expression. Since $\partial\Pi_i/\partial q_\ell = \theta_i\pi_n\beta_i Z_\ell \geq 0$ and $\partial\Pi_i/\partial Z_\ell = -\theta_i\pi_n\beta_i(1 - q_\ell) \leq 0$, the same signs hold for $\mathbb{E}[\Pi_i]$. $\square$

Proposition 2 (Intensity of use). *Among participating households with intensity β and the same (h, y) , for which I write F_ℓ and f_ℓ without the conditioning arguments, the probability of investing, $1 - F_\ell(a(s)\pi_n\beta\Delta_\ell)$, and expected willingness to pay, $\omega(\mathbb{E}[x_i | h, y] - a(s)\pi_n\beta\Delta_\ell)$, decrease with β at any s . The effect of salience on expected willingness to pay, $-\omega(a_1 - a_0)\pi_n\beta\Delta_\ell$, increases in absolute value with β . The effect on the probability of investing, $F_\ell(a_0\pi_n\beta\Delta_\ell) - F_\ell(a_1\pi_n\beta\Delta_\ell)$, can rise or fall with β .*

Proof. The first two expressions decrease with β because F_ℓ is nondecreasing and W_i is linear in β . The third is linear in β with slope $-\omega(a_1 - a_0)\pi_n\Delta_\ell \leq 0$. For the last, let $m \equiv \pi_n\beta\Delta_\ell$; the derivative of $F_\ell(a_0m) - F_\ell(a_1m)$ with respect to m is $a_0f_\ell(a_0m) - a_1f_\ell(a_1m)$, which is negative if the density is roughly flat between a_0m and a_1m and positive if it falls steeply over that range. $\square$

Proposition 3 (Who responds). *A household responds to salience only if $\Pi_i > 0$ and $a_0\Pi_i \leq x_i < a_1\Pi_i$. Households that cannot afford to participate, $c(y_i) > \beta_i\Delta_\ell$, and households whose net return exceeds $a_1\Pi_i$ do not respond.*

Proof. The first statement restates the interval in Proposition 1. The poorest households have $\theta_i = 0$ by step 2, so $\Pi_i = 0$. Households with $x_i \geq a_1\Pi_i$ invest in both arms. Since C_i decreases with y_i , households with more resources have larger x_i and are more likely to be in this group. The response is therefore concentrated among participating households with intermediate net returns, rather than among the poorest. $\square$

Proposition 4 (The gap conditional on parental human capital). *Suppose the return ρ is common, and the distribution of C_i given (h_i, y_i) , and hence the distribution function F of x_i , is the same on both sides. If, conditional on (h, y) , Π_i on the east side first-order stochastically dominates Π_i on the west side, attendance conditional on (h, y) is weakly lower on the east side for any s . The gap is zero for children with $\pi_n = 0$, and it weakly increases with π_n if $z \mapsto zf(z)$ is nondecreasing on the relevant range or if no west-side household participates.*

Proof. Conditional on (h, y) , attendance on side ℓ is $1 - \mathbb{E}_\ell[F(a(s)\Pi_i)]$. Because F is nondecreasing, first-order dominance of Π_i on the east side implies $\mathbb{E}_E[F(a(s)\Pi_i)] \geq \mathbb{E}_W[F(a(s)\Pi_i)]$, so attendance is weakly lower on the east side. For the age statement, write $X_i \equiv a(s)\theta_i\beta_i\Delta_\ell$, so that $a(s)\Pi_i = \pi_n X_i$ and the gap is $\mathbb{E}_E[F(\pi_n X_i)] - \mathbb{E}_W[F(\pi_n X_i)]$. At $\pi_n = 0$, both terms equal $F(0)$ and the gap is zero. Its derivative with respect to π_n is $\mathbb{E}_E[X_i f(\pi_n X_i)] - \mathbb{E}_W[X_i f(\pi_n X_i)]$. If $zf(z)$ is nondecreasing, then $Xf(\pi_n X) = \pi_n^{-1}(\pi_n X)f(\pi_n X)$ is nondecreasing in X , and first-order dominance makes the derivative nonnegative. If no west-side household participates, $X_i = 0$ on the west side and the second expectation is zero. $\square$

Equalizing provision equalizes the distribution of C_i across sides but leaves Π_i unchanged, so the gap in Proposition 4 remains.

Proposition 5 (Required return). *The return at which household i is indifferent about investing is $\bar{\rho}_i = C_i + a(s)\Pi_i$, which is higher for households with a larger Π_i at the same cost and the same perceived return.*

Proof. Setting $\rho - C_i = a(s)\Pi_i$ in the investment condition and solving for ρ gives $\bar{\rho}_i$, which increases with Π_i . $\square$

F.3 The Dynamic Model

The dynamic model follows Bisin and Verdier (2001). It adds steps 4 and 5 of the timing to the static model and asks how the participating share evolves across generations. To simplify notation, I drop the subscript ℓ in this subsection, and I abstract from the participation cost, so that every household disposed to participate does so and q_t is both the disposed and the participating share.

Step 1: How a child acquires a disposition. A child acquires a disposition in two stages. First, the parent tries to transmit its own disposition and succeeds with probability τ , chosen by the parent. If this direct transmission fails, the child adopts the disposition of an adult drawn at random from the neighborhood, who is disposed to participate with probability q_t . Denote by P^{jk} the probability that the child of a parent with disposition j ends up with disposition k . For a parent disposed to participate ($j = 1$), the child participates if direct transmission succeeds, or if it fails and the random adult participates:

$$P^{11} = \tau_1 + (1 - \tau_1) q_t, \quad P^{10} = (1 - \tau_1)(1 - q_t).$$

For a parent not disposed to participate ($j = 0$):

$$P^{00} = \tau_0 + (1 - \tau_0)(1 - q_t), \quad P^{01} = (1 - \tau_0) q_t.$$

In each case, the two probabilities add up to one.

Step 2: What parents want. Parents evaluate their children's dispositions with their own preferences, as in [Bisin and Verdier \(2001\)](#). A parent disposed to participate values a child who participates over one who does not by

$$V^1 = \lambda \max\{\Delta(q_t, Z), 0\},$$

the value of keeping access to local exchange, with $\lambda > 0$. When local exchange has positive value, $V^1 = \lambda \Delta(q_t, Z)$; when it has none, the parent gains nothing from transmitting participation. A parent not disposed to participate values a child who relies on outside channels over one who participates by $V^0 = b(Z) > 0$, with $b'(Z) \geq 0$. Parents evaluate these gains at the current participating share q_t . This is the standard simplification in [Bisin and Verdier \(2001\)](#), and it does not affect the rest points, where $q_{t+1} = q_t$.

Step 3: Optimal socialization. A parent disposed to participate chooses τ_1 to maximize the expected value of its child's disposition minus the cost of socialization. Normalizing the value of a non-participating child to zero, the problem is

$$\max_{\tau_1 \in [0,1]} P^{11}(\tau_1) V^1 - \frac{\tau_1^2}{2k} = \max_{\tau_1 \in [0,1]} [q_t + \tau_1(1 - q_t)] V^1 - \frac{\tau_1^2}{2k}.$$

The first-order condition is $(1 - q_t)V^1 - \tau_1/k = 0$, and the second derivative is $-1/k < 0$, so

$$\tau_1 = k(1 - q_t)V^1.$$

By the same steps, a parent not disposed to participate chooses

$$\tau_0 = k q_t V^0.$$

Both solutions lie in $[0, 1]$ if $k \max\{\lambda R, b(Z)\} \leq 1$, because $V^1 \leq \lambda R$. Each parent socializes more when its own disposition is scarce, because the random adult is then less likely to transmit it. This is the cultural substitution of [Bisin and Verdier \(2001\)](#). A parent disposed to participate also socializes more when local exchange is more valuable.

Step 4: The law of motion. The participating share in generation $t + 1$ is the share of children who acquire the disposition:

$$q_{t+1} = q_t P^{11} + (1 - q_t) P^{01} = q_t [\tau_1 + (1 - \tau_1)q_t] + (1 - q_t)(1 - \tau_0) q_t.$$

Expanding the right-hand side,

$$q_{t+1} = q_t \tau_1 + q_t^2 - q_t^2 \tau_1 + q_t - q_t^2 - q_t \tau_0 + q_t^2 \tau_0 = q_t + q_t(1 - q_t)(\tau_1 - \tau_0).$$

Substituting the optimal efforts from step 3,

$$q_{t+1} - q_t = k q_t(1 - q_t) \Phi(q_t, Z), \quad \Phi(q, Z) = (1 - q) \lambda \max\{\Delta(q, Z), 0\} - q b(Z),$$

which is (6) when local exchange has positive value. The share rises when $\Phi > 0$, that is, when the incentive of participating parents to transmit, weighted by the share of non-participants, exceeds the opposing incentive of non-participating parents, weighted by the share of participants.

Step 5: Rest points and their stability. Write the law of motion as $q_{t+1} = g(q_t)$, with $g(q) = q + kq(1 - q)\Phi(q, Z)$. A rest point satisfies $g(q) = q$, so it is either $q = 0$, $q = 1$, or an interior zero of $\Phi(\cdot, Z)$. At an interior zero $\hat{q}$,

$$g'(\hat{q}) = 1 + k\hat{q}(1 - \hat{q}) \frac{\partial \Phi}{\partial q}(\hat{q}, Z).$$

The rest point is locally stable if $|g'(\hat{q})| < 1$. This holds when $\partial \Phi / \partial q < 0$ at $\hat{q}$, that is, when Φ crosses zero from above, and k is small enough that $k\hat{q}(1 - \hat{q})|\partial \Phi / \partial q| < 2$. If, in addition, $k\hat{q}(1 - \hat{q})|\partial \Phi / \partial q| < 1$, convergence is monotone. A zero where Φ crosses from below has $g'(\hat{q}) > 1$ and is unstable. I assume throughout that k is small enough, so stability is determined by the sign of $\partial \Phi / \partial q$.

Step 6: The linear case. With $\Delta(q, Z) = R - Z(1 - q)$, let $u \equiv 1 - q$ denote the share not participating. Where $\Delta > 0$,

$$\Phi = \lambda u (R - Zu) - (1 - u)b = -\lambda Z u^2 + (\lambda R + b)u - b,$$

a concave quadratic in u . Its values at the boundaries are $\Phi = -b < 0$ at $u = 0$ ($q = 1$) and $\Phi = \lambda(R - Z)$ at $u = 1$ ($q = 0$). Its roots are

$$u_{\pm} = \frac{(\lambda R + b) \pm \sqrt{\mathcal{D}}}{2\lambda Z}, \quad \mathcal{D} \equiv (\lambda R + b)^2 - 4\lambda Z b.$$

Where $\Delta \leq 0$, $\Phi = -qb < 0$. At any interior zero, $(1 - q)\lambda\Delta = qb > 0$, so $\Delta > 0$: every interior zero of Φ is a root of the quadratic.

Proposition 6 (Two regimes). *Let $\Delta(q, Z) = R - Z(1 - q)$ and let k be small enough.*

1. *If $R > Z$, there is a unique interior rest point $q^* = 1 - u_-$, which is stable and attracts every interior initial share. It decreases with Z .*
2. *If $R < Z$ and either $\mathcal{D} < 0$ or the vertex $(\lambda R + b)/(2\lambda Z)$ lies outside $(0, 1)$, there is no interior rest point, and the share declines toward zero from any initial level.*
3. *If $R < Z$, $\mathcal{D} > 0$, and the vertex lies in $(0, 1)$, there are two interior rest points, $q_c = 1 - u_+$ and $q^\dagger = 1 - u_-$, with $q_c < q^\dagger$. The rest points $q = 0$ and $q^\dagger$ are stable and q_c is unstable: shares above q_c converge to $q^\dagger$, and shares below it decline toward zero. An increase in Z raises q_c and lowers $q^\dagger$, and for Z large enough the two rest points disappear.*

Proof. (1) If $R > Z$, then $\Delta > 0$ for all q , and Φ is negative at $u = 0$ and positive at $u = 1$. A concave function with these boundary values crosses zero exactly once on $(0, 1)$: the smaller root u_- lies in $(0, 1)$, and the larger root u_+ exceeds one. In terms of q , Φ is positive below $q^* = 1 - u_-$ and negative above it, so Φ crosses zero from above, and by step 5 the rest point is stable. Since $\Phi > 0$ below q^* and $\Phi < 0$ above it, every interior initial share moves toward q^* .

For the comparative static, $\partial\Phi/\partial Z = -\lambda u^2 - qb'(Z) < 0$ and $\partial\Phi/\partial q < 0$ at q^* , so the implicit function theorem gives $dq^*/dZ = -(\partial\Phi/\partial Z)/(\partial\Phi/\partial q) < 0$.

(2) If $R < Z$, Φ is negative at both $u = 0$ and $u = 1$. A concave quadratic that is negative at both ends of $(0, 1)$ is positive somewhere in between only if both of its roots lie in $(0, 1)$, which requires $\mathcal{D} > 0$ and its vertex in $(0, 1)$. Otherwise $\Phi < 0$ on $(0, 1)$, so $q_{t+1} < q_t$ at every interior share.

(3) Under the stated conditions, both roots lie in $(0, 1)$ and $\Phi > 0$ between them. In terms of q , $\Phi < 0$ below $q_c = 1 - u_+$, $\Phi > 0$ between q_c and $q^\dagger = 1 - u_-$, and $\Phi < 0$ above $q^\dagger$. Φ therefore crosses zero from below at q_c , which is unstable, and from above at $q^\dagger$, which is stable; $q = 0$ is stable because $\Phi < 0$ just above it. Because $\partial\Phi/\partial Z < 0$, an increase in Z shifts Φ down, so the interval where $\Phi > 0$ shrinks from both ends, which raises q_c and lowers $q^\dagger$, until the maximum of Φ falls below zero and both rest points disappear. $\square$

Worked example. The parameters of Figure 7, with $\lambda = k = 1$, illustrate the formulas. In panel (a), $R = 0.6$, $Z = 0.2$, and $b = 0.1$, so $\lambda R + b = 0.7$, $\mathcal{D} = 0.49 - 0.08 = 0.41$, and $u_- = (0.7 - 0.640)/0.4 \approx 0.149$, which gives $q^* \approx 0.85$. With $Z = 0.4$ and $b = 0.2$, the same steps give $\mathcal{D} = 0.64 - 0.32 = 0.32$, $u_- = (0.8 - 0.566)/0.8 \approx 0.293$, and $q^* \approx 0.71$. In panel (b), $R = 0.4$, $Z = 0.6$, and $b = 0.03$, so $\lambda R + b = 0.43$ and $\mathcal{D} = 0.1849 - 0.072 = 0.1129$. The roots are $u_+ = (0.43 + 0.336)/1.2 \approx 0.638$ and $u_- = (0.43 - 0.336)/1.2 \approx 0.078$, which give $q_c \approx 0.36$ and $q^\dagger \approx 0.92$.

Step 7: How fast participation adjusts. Near a stable interior rest point $\hat{q}$, the distance to the rest point shrinks by the factor $g'(\hat{q})$ each generation. The number of generations needed to halve that distance is approximately

$$\frac{\ln 2}{k \hat{q}(1 - \hat{q}) |\partial\Phi/\partial q(\hat{q}, Z)|}.$$

Adjustment is slow when socialization responds little to incentives (small k), when the rest point is close to zero or one, or when Φ is flat around it. After a permanent improvement in outside channels from Z_0 to Z_1 in the first regime, the share moves from $q^*(Z_0)$ toward the lower level $q^*(Z_1)$ at this rate. In the third case of Proposition 6, a neighborhood with $q_t > q_c(Z_1)$ when outside channels improve converges to $q^\dagger(Z_1)$ instead of declining toward zero.

Step 8: From participation to the gap.

Lemma 1 (Participation and the gap). *Suppose local exchange has nonnegative value on both sides, $\Delta_{\ell t} \geq 0$, all households in a neighborhood have intensity β and can afford to participate, that x_i is independent of the disposition to participate given parental human capital h , and let F_h be the distribution function of x_i given h , common to both sides. At weight a_0 , attendance among*

children aged n in neighborhood ℓ is

$$A_{\ell t}(h) = 1 - F_h(0) - q_{\ell t} [F_h(a_0 \pi_n \beta \Delta_{\ell t}) - F_h(0)],$$

and the gap $A_{Wt}(h) - A_{Et}(h)$ is (7). The gap increases with q_{Et} and with Δ_{Et} , and it is zero when both sides have the same $q_{\ell t}$ and $\Delta_{\ell t}$.

Proof. A share $1 - q_{\ell t}$ of households does not participate and invests if $x_i \geq 0$, which happens with probability $1 - F_h(0)$. A share $q_{\ell t}$ participates and invests if $x_i \geq a_0 \pi_n \beta \Delta_{\ell t}$, which happens with probability $1 - F_h(a_0 \pi_n \beta \Delta_{\ell t})$. Attendance is the weighted average,

$$A_{\ell t}(h) = (1 - q_{\ell t}) [1 - F_h(0)] + q_{\ell t} [1 - F_h(a_0 \pi_n \beta \Delta_{\ell t})],$$

which rearranges to the expression stated. Subtracting the east-side expression from the west-side one gives (7). The bracket multiplying q_{Et} is nonnegative because $\Delta_{Et} \geq 0$ and F_h is nondecreasing, so the gap increases with q_{Et} , and it increases with Δ_{Et} for the same reason. $\square$

Step 9: Human capital. The child's education becomes the next generation's parental human capital, $h_{i,t+1} = e_{i,t}$. Because C decreases with h , lower investment among participating households raises the next generation's cost of education, which lowers attendance beyond Lemma 1. This channel operates through h , so conditioning on parental human capital removes it and leaves the gap in Lemma 1. As generations pass, more of the original gap is embodied in parental human capital, so conditioning on it absorbs a larger share of the raw gap.

Step 10: Residential turnover. With turnover, a share μ of adults is replaced each generation by entrants whose participating share is q^M , so the law of motion becomes

$$q_{t+1} = (1 - \mu) g(q_t) + \mu q^M,$$

which is (8). A one-time replacement of residents changes q_t once and then lets g act. In the first regime of Proposition 6, the share returns toward q^* . In the third case, a one-time change that moves the share from above q_c to below it sends the neighborhood toward zero, while a smaller change is followed by a return toward $q^\dagger$. A permanent inflow of entrants with $q^M < q^\dagger$ lowers the rest point of the modified law of motion and, if $q^M < q_c$ and μ is large enough, eliminates the upper rest point.

Appendix References

- 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 17).
- 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. 30).
- 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. 9).
- 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. 15 and 21).
- 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. 28 and 29).
- Bagagli, Sara**, “The (Express)Way to Segregation: Evidence from Chicag,” Technical Report 2025. (cited on p. 9).
- 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 p. 137).
- 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 18).
- Borusyak, Kirill and Peter Hull**, “Nonrandom Exposure to Exogenous Shocks,” *Econometrica*, 11 2023, 91 (6), 2155–2185. (cited on p. 20).
- Buringh, E.**, “Urbanisation hub: The clio-infra database on urban settlement sizes: 1500–2000.” 2018. (cited on p. 8).
- Bursztyn, Leonardo and David Y. Yang**, “Misperceptions About Others,” *Annual Review of Economics*, 8 2022, 14, 425–452.

- 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 and 14).
- Cappelen, Alexander W., Benjamin Enke, and Bertil Tungodden**, “Universalism: Global Evidence,” *American Economic Review*, 1 2025, 115, 43–76. (cited on p. 28).
- 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. 11, 15, and 16).
- Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebe, 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.
- 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 p. 9).
- Cohn, Alain and Michel André Maréchal**, “Priming in Economics,” *Current Opinion in Psychology*, 2016, 12, 17–21.
- Cunha, Flavio and James Heckman**, “The Technology of Skill Formation,” *American Economic Review*, 2007, 97 (2), 31–47. (cited on pp. 2, 17, and 33).
- de Arregui, Domingo Lázaro**, *Descripción de la Nueva Galicia* 1946. (cited on p. 8).
- Enke, Benjamin, 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. 28).
- Fryer, Roland G. and Steven D. Levitt**, “The Causes and Consequences of Distinctively Black Names,” *The Quarterly Journal of Economics*, 2004, 119 (3), 767–805. (cited on p. 28).

- Galor, Oded**, “Inequality, Human Capital Formation, and the Process of Development,” in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, Vol. 4, Elsevier, 2011, chapter 5, pp. 441–493. (cited on pp. 2 and 17).
- **and Daniel Tsiddon**, “The Distribution of Human Capital and Economic Growth,” *Journal of Economic Growth*, 1997, 2 (1), 93–124.
- **and Joseph Zeira**, “Income Distribution and Macroeconomics,” *The Review of Economic Studies*, 1993, 60 (1), 35–52. (cited on pp. 2 and 17).
- 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. 8).
- 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 p. 23).
- 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 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.
- 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. 9).