

EDUCATIONAL DECISION-MAKING IN CONTEXTS OF INEQUALITY: A DATA ANALYSIS OF THE GAP BETWEEN POVERTY AND ACADEMIC SUCCESS

Fredy Olarte, & Liz Karen Herrera
Universidad Nacional de Colombia (Colombia)

Abstract

The expansion of online learning in emerging economies has intensified interest in the structural determinants of student persistence. This study investigates how district-level socioeconomic configurations relate to academic risk among 3,678 students enrolled in a virtual technology program in Bogotá, Colombia. Using a cross-sectional design that integrates academic records with district-level monetary poverty rates and Gini coefficients across 19 localities, weighted Pearson correlations reveal a counterintuitive paradox: monetary poverty shows a weak, non-significant correlation with academic risk ($r = 0.157$; $p = 0.53$), whereas the local Gini coefficient exhibits a strong negative correlation ($r = -0.647$; $p < 0.01$). Low-income students in high-inequality, socioeconomically mixed localities show success rates approximately three times higher than comparable peers in homogeneously poor peripheral localities. Drawing on theories of urban opportunity structures, social capital, and the digital divide, we interpret the Gini coefficient as a proxy for access to opportunity infrastructure. Findings suggest that territorial interventions targeting digital infrastructure may be more effective for student retention than individual-level income support alone.

Keywords: Educational data mining, urban inequality, student persistence, digital divide, evidence-based policy.

1. Introduction

The global expansion of virtual and hybrid education has opened pathways for populations historically excluded from formal learning. In Colombia, virtual technological programs have been promoted as a mechanism to broaden access and reduce persistent inequalities in human capital formation (MEN, 2022; UNESCO, 2021). However, expanded access does not automatically translate into academic success: retention remains a critical challenge, particularly among economically precarious students (Tinto, 1987; Bean & Metzner, 1985).

Educational policy has long emphasized household income as the primary predictor of dropout risk, guiding interventions toward scholarships and emergency economic support. While legitimate, this approach may overlook a more complex constellation of structural factors rooted in the urban environment. Bogotá's 19 administrative localities exhibit profound socioeconomic heterogeneity, ranging from consolidated urban core zones with dense service infrastructure to peripheral localities characterized by concentrated poverty and limited public investment (Castells, 1996; Secretaría Distrital de Planeación, 2022).

This study examines whether district-level socioeconomic indicators—monetary poverty rates and income inequality as measured by the Gini coefficient—predict academic persistence in a virtual program, challenging the assumption that individual-level financial hardship is the dominant territorial predictor of student dropout.

2. Theoretical framework

This study draws on four interconnected bodies of literature. First, student persistence theory highlights that structural and contextual variables—not only individual motivation—condition academic retention (Tinto, 1987; Bean & Metzner, 1985; Romero & Ventura, 2020). Second, urban opportunity structures theory posits that residential location shapes consequential access to resources: neighborhoods of concentrated poverty limit access to quality infrastructure, social networks, and professional role

models—a condition of "double disadvantage" (Wilson, 1987; Massey & Denton, 1993)—while socioeconomically heterogeneous environments may offer protective effects through social mixing (Musterd & Andersson, 2005; Atkinson & Kintrea, 2000).

Third, digital divide scholarship has evolved beyond binary access to encompass quality of connectivity and the availability of complementary public digital infrastructure (Van Dijk, 2006; Warschauer, 2003). Peripheral urban localities suffer a "digital desert" effect: even if households possess basic internet access, the density of public Wi-Fi, libraries, and community digital spaces is substantially lower than in economically dynamic central areas (Graham & Marvin, 2001; CEPAL, 2021). For students in fully virtual programs, this spatial inequality in infrastructure may be decisive. Fourth, agglomeration effects and social capital theory suggest that economically diverse localities generate positive externalities including bridging social capital (Putnam, 2000), knowledge spillovers, and shared infrastructure (Glaeser, 2011) that may partially offset material disadvantages for low-income residents.

3. Methodology

This study employs a cross-sectional correlational design (Creswell & Creswell, 2018). Administrative academic records of 3,678 students enrolled in a public virtual technology program in Bogotá (2022–2023) were integrated with district-level socioeconomic indicators from DANE and the Secretaría Distrital de Planeación. The unit of analysis is the locality; student-level data are aggregated to produce district-level academic risk rates.

The dependent variable is the academic risk rate per locality (proportion of students who failed or withdrew). Independent variables are: (1) district monetary poverty rate (percentage below the poverty line, per DANE); and (2) district Gini coefficient (income inequality within the locality). Socioeconomic stratum (1–6) was used for subgroup analysis. Weighted Pearson correlation coefficients were computed using locality population share as weight, with sensitivity analyses excluding localities with minimal samples (Sumapaz, $n = 2$; La Candelaria, $n = 7$) to ensure statistical reliability.

4. Results

4.1. Descriptive overview

Table 1 presents the distribution of students, academic risk rates, and socioeconomic indicators for the 12 most representative localities ($n \geq 50$), covering 93% of total enrollment. Academic risk rates range from 25.5% (Santa Fe) to 57.9% (Rafael Uribe Uribe). A visual inspection suggests that localities with higher Gini coefficients—Santa Fe (0.65), Usaquén (0.54), Suba (0.56)—tend to exhibit lower risk rates, while low-Gini peripheral localities—Usme (0.42), Ciudad Bolívar (0.47)—show higher rates. No equivalent directional pattern is apparent for monetary poverty.

Table 1. Distribution of Students, Academic Risk, and Socioeconomic Indicators by Locality ($n \geq 50$).

Locality	n	Risk %	Gini	Poverty %
Santa Fe	51	25.5	0.65	47.7
Usaquén	137	32.8	0.54	16.0
Fontibón	191	40.8	0.49	19.8
San Cristóbal	248	43.5	0.45	48.5
Bosa	473	47.1	0.46	53.2
Engativá	486	47.5	0.47	24.9
Kennedy	612	47.7	0.50	37.0
Suba	584	50.9	0.56	25.8
Tunjuelito	102	51.0	0.48	40.2
Usme	231	52.4	0.42	57.8
Ciudad Bolívar	359	54.0	0.47	57.4
Rafael Uribe	190	57.9	0.50	50.0

Note. Sorted by academic risk rate (ascending). Gini: district income inequality coefficient. Poverty %: population below the monetary poverty line. Source: DANE & Secretaría Distrital de Planeación (2022); institutional academic records.

4.2. Correlation analysis

The weighted Pearson correlation between district monetary poverty rate and academic risk was $r = 0.157$ ($p = 0.53$)—weak and statistically non-significant, consistent across sensitivity analyses. By contrast, the correlation between the district Gini coefficient and academic risk was $r = -0.647$ ($p < 0.01$)—strong, negative, and statistically significant. Localities with greater income heterogeneity are associated with substantially lower academic risk rates. Subgroup analysis of strata 1–2 students (62% of the sample) confirms that this relationship is most pronounced among low-income students: a student in stratum 1 or 2 residing in a high-Gini, socioeconomically mixed locality shows an academic success rate approximately three times higher than a comparable peer in a low-Gini, homogeneously poor peripheral locality.

5. Discussion

The absence of a significant correlation between monetary poverty and academic risk challenges the prevailing policy assumption that aggregate poverty rates are reliable territorial predictors of educational disadvantage. The strong negative Gini-risk correlation, by contrast, supports a reading of urban inequality that foregrounds socioeconomic heterogeneity as a structural resource rather than a liability.

We interpret this pattern through the concept of urban opportunity infrastructure (Wilson, 1987; Massey & Denton, 1993). In Bogotá, high-Gini localities—such as Santa Fe, Usaquén, and Suba—are typically more central and functionally diverse, with greater density of public libraries, transportation networks, and digital access points. Even low-income students in these localities may benefit from proximity to resources that support virtual learning: stable connectivity, physical study spaces, and community digital infrastructure. This is consistent with the "digital desert" framework (Warschauer, 2003; Graham & Marvin, 2001): peripheral, low-Gini localities in Bogotá are not only economically disadvantaged but informationally isolated—a condition that compounds constraints on academic engagement beyond what household income alone can explain.

The social capital dimension is equally relevant. Drawing on Putnam (2000), socioeconomically diverse localities may facilitate bridging social capital networks that provide low-income students access to varied information, informal mentoring, and peer learning unavailable in more homogeneous environments. Agglomeration effects in economically dense districts may additionally generate educational externalities that benefit all residents (Glaeser, 2011). Together, these mechanisms suggest that the Gini coefficient functions as an indirect indicator of the socioeconomic complexity and infrastructural density of the urban environment—what we term the "infrastructure of opportunity."

Several limitations warrant acknowledgment. The ecological correlation design precludes individual-level causal inference, and the cross-sectional approach limits temporal conclusions. The Gini coefficient may capture sources of variation beyond income heterogeneity—including urban primacy and political economy of public investment—that future research should disentangle.

6. Conclusions

This study provides empirical evidence that district-level income inequality is a stronger predictor of academic persistence than aggregate poverty rates among students in a virtual technology program. These findings have concrete implications for educational policy. Individual-level economic interventions, while necessary, are insufficient as the sole retention mechanism. Complementary territorial strategies should target digital infrastructure investment in peripheral localities, the expansion of community learning spaces with reliable connectivity, and the integration of educational support into neighborhoods affected by spatial isolation and digital scarcity.

More broadly, findings support the incorporation of district-level Gini coefficients and poverty maps into institutional early warning systems, and advocate for cross-sectoral coordination between education and urban development policy. Addressing the "digital desert" effect in Bogotá's peripheral localities requires not only educational investment but urban infrastructure planning that ensures equitable distribution of digital resources. These conclusions apply, with contextual adaptation, to other cities in Latin America and emerging economies where virtual education expansion is occurring against a backdrop of uneven urban development.

References

- Atkinson, R., & Kintrea, K. (2000). Owner-occupation, social mix and neighbourhood impacts. *Policy and Politics*, 28(1), 93–108.
- Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, 55(4), 485–540.
- Castells, M. (1996). *The rise of the network society*. Blackwell.
- CEPAL. (2021). *Tecnologías digitales para un nuevo futuro*. Naciones Unidas.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
- DANE. (2023). *Encuesta multipropósito Bogotá 2021: Principales resultados*. DANE.
- Glaeser, E. L. (2011). *Triumph of the city*. Penguin Press.
- Graham, S., & Marvin, S. (2001). *Splintering urbanism*. Routledge.
- Massey, D. S., & Denton, N. A. (1993). *American apartheid*. Harvard University Press.
- MEN (Ministerio de Educación Nacional). (2022). *Informe de gestión 2022*. MEN Colombia.
- Musterd, S., & Andersson, R. (2005). Housing mix, social mix, and social opportunities. *Urban Affairs Review*, 40(6), 761–790.
- Putnam, R. D. (2000). *Bowling alone*. Simon & Schuster.
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1355.
- Secretaría Distrital de Planeación de Bogotá. (2022). *Encuesta multipropósito 2021: Resultados por localidad*. Alcaldía Mayor de Bogotá.
- Tinto, V. (1987). *Leaving college*. University of Chicago Press.
- UNESCO. (2021). *Rethinking education: Towards a global common good?* UNESCO Publishing.
- Van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics*, 34(4–5), 221–235.
- Warschauer, M. (2003). *Technology and social inclusion: Rethinking the digital divide*. MIT Press.
- Wilson, W. J. (1987). *The truly disadvantaged*. University of Chicago Press.