I Cover Letter and References

Organizing Committee
Workshop on Development Economics and Networks for Early-Career Researchers
National Bureau of Economic Research

Dear Professors Breza, Viviano, and Chandrasekhar,

I am writing to apply to the Workshop on Development Economics and Networks for Early-Career Researchers. I am a Ph.D. candidate in Economics and Education at Teachers College, Columbia University, where my research focuses on the intersection of early childhood development, neuroeconomics, and development economics. My dissertation explores how testing environments affect scores heterogeneously by SES background, with consequential downstream effects on college access. I would welcome the opportunity to present my work and learn from the frontier of empirical network research alongside the workshop's participants.

My broader research agenda examines how measurement shapes opportunity in developing-country and conflict-affected settings, including work on the AI shock and the signal value of grades, an intervention on skill development in Morocco's Pioneer Schools, and ongoing work validating measures of academic, social, and emotional outcomes for conflict-affected children in the Democratic Republic of Congo. Across these projects, I engage closely with questions of information, measurement, and behavior in resource-constrained settings. I am eager to get feedback and build connections with the broader community of early-career development economists that this workshop convenes.

Please find attached my curriculum vitae, a statement of motivation and research interests, and a working paper for consideration. I have also asked my references to provide a letter of recommendation should I be shortlisted for participation.

Thank you for considering my application. I look forward to the possibility of joining the workshop.

Sincerely,

Alejandra Campos Alejandra Campos
311 W 111th St, NY, NY, 10027
[202-5150276]
aac2271@tc.columbia.edu

II References

Alex Eble
Associate Professor of Economics and Education
Department of Economics
Teachers College, Columbia University
Email: eble@tc.columbia.edu

Jonah Rockoff
Paul Garret Professor
Columbia Business School
Columbia University
Email: jonah.rockoff@columbia.edu

III Statement

I was born and raised in Colombia, a country where it is estimated to take 250 years for a family in extreme poverty to reach the middle class. Growing up in this context, I witnessed firsthand how deeply these disparities influence life opportunities, particularly in education. These disparities were present within my own family: my father began working at the age of twelve to support his mother and siblings, yet through perseverance and determination, he earned a degree from a public university.

This personal history evolved into a scholarly inquiry: what differentiated my father from his siblings, who were unable to pursue further schooling? It was not a lack of intelligence, but rather a set of capacities—perseverance, creativity, focus—that helped him navigate adversity. Yet these abilities, while essential to his success, were largely imperceptible to conventional measures of merit or academic potential. These observations now drive my research into how early-life adversity shapes long-term human potential, which skills emerge in response to scarcity, and how education systems can better identify and support exceptionally capable students whose skills fall outside standard measures.

My current research, as a Ph.D. candidate in Economics and Education at Columbia University, is situated at the intersection of development economics, skill measurement, and neuroeconomics. Drawing on theories of adaptive skill formation, my dissertation examines how students of different socio-economic backgrounds perform under unfavorable environments. Specifically, I investigate whether exposure to instability fosters context-specific skills—such as flexible problem-solving, rapid attention shifting, and persistence under stress—that may be relevant but not fully captured by standard tests. By focusing on when and where skills are expressed, rather than only average performance, my work incorporates environmental context as an additional dimension in the measurement of ability. At this stage of my Ph.D., my project would particularly benefit from the structured feedback and mentorship of development economists —precisely the environment fostered by the NBER Workshop.

My training is complemented by extensive policy and field experience. At the Inter-American Development Bank, I analyzed gaps in education policies across Latin America, gaining firsthand insight into how structural and economic disadvantages constrain access to high-quality learning opportunities. Later, as a fellow at the University of Chicago, I studied early childhood programs designed to foster skill development in contexts of adversity. I have also completed formal coursework in psychometrics and currently serve as a measurement lead on a large-scale project in Morocco, where I work on item design, psychometric validation, and the linking of national assessments to international benchmarks such as PISA and TIMSS. Together, these experiences have reinforced my interest in how assessment design influences the identification of high-performing students and their access to advanced learning pathways.

These experiences collectively shape both the questions I ask and the scholarly communities I strive to join. At the NBER workshop, I am particularly eager to engage with the development economics community. NBER's unique emphasis on mentorship, inclusion, and rigorous feedback provides the ideal environment to refine this project as I prepare for the 2027–28 job market. Ultimately, as one of the few Colombian women pursuing a Ph.D. abroad, I am committed to using my academic training to build fairer, more inclusive education systems—ones that recognize diverse forms of ability and empower students from marginalized backgrounds to realize their full potential. Thank you for your time and consideration.

Sincerely,

Alejandra Campos

Alejandra Campos Quintero

+1 (202)-5150276 – aac2271@tc.columbia.edu – Linkedin

EDUCATION

Ph.D. Economics and Education - Columbia University 2022 – 2027 (Expected)
New York, NY | Advisor: Alex Eble, Jonah Rockoff (GPA: 3.9/4)

Pre- Doctoral Fellow - University of Chicago 2020 – 2022
Chicago, IL, USA| Advisor: Professor James Heckman (Nobel Laureate)

M.A. in Economics- Universidad de los Andes 2015– 2017
(GPA: 4/4)

B.A. in Economics- Universidad de los Andes 2011– 2015
(GPA: 3.9/4)

RESEARCH FIELDS

Economics of Education, Development Economics, Psychometrics, Labor Economics, Behavioral Economics.

PUBLICATIONS

"Interactions: Do Teacher Behaviors Predict Achievement, Executive Function, and Non-Cognitive Outcomes in Elementary School?" - with Schady, Norbert & Carneiro, Pedro & Cruz-Aguayo, Yyannú . Revised and Resubmitted to **The Journal of Political Economy (JPE**) April 2026.

"Developing socioemotional skills in early adolescence: Experimental evidence from Morocco's public schools" - with de Barros, Glewwe, P., Kumar, N., & Lépine, L. Manuscript conditionally accepted for publication in the **Journal of Development Economics** (2025)

"The reliability of classroom observations and student surveys in non-research settings: Evidence from Argentina" – with Ganimian, Alejandro J., and Andrew D. Ho (2025). Accepted at the **Education Evaluation and Policy Analysis**. (**EdWorkingPaper**: 25-1322). https://doi.org/10.26300/nvmr-5e94

Master´s Thesis: "Roads to Education: Impact of Road Infrastructure on Education (2005 – 2015)". **Documento CEDE** No. 2017-04, Department of Economics, University of the Andes, January 2017. (4.8/5). http://dx.doi.org/10.2139/ssrn.2921472

WORKING PAPERS

- **JMP:** "Hidden Strengths: Adversity, Adaptation, and the Measurement of Skills." - 2026
- Beyond Basics: Whole School Reform and Early Adolescent Development - with Andreas de Barros, Najiba El Amrani Mida, Paul Glewwe, Nikhil Kumar, Laure Lépine
- Paraprofessional Experience and Teacher Quality: Evidence from New York City Public Schools – with Jonah Rockoff
- AI adoption and the value of academic credentials– with Thea Dowrich

PROFESSIONAL EXPERIENCE

Justice Rising. **Jul 2026 -Present**
Measurement Advisor

Benchmarked a school-quality assessment tool against six leading international education-quality frameworks to identify measurement gaps and align it with established best practices for education in conflict settings.

Morocco Innovation and Evaluation Lab, J-PAL. **Jun 2024- Present**
Measurement and Psychometrics Consultant

Led assessment development and psychometric analysis for large-scale education evaluations with Morocco's Ministry of Education, including item development, IRT-based scaling and linking, and analysis of student learning outcomes in middle schools.

Produce knowledge products linking national assessment results to PISA/TIMSS international benchmarks, directly informing ministry policy decisions.

Designed and implemented evaluation framework to assess learning-tool quality, translating results into recommendations.

Columbia Business School - Columbia University. **Jan 2025 – Present**
Graduate Research Assistant - Professor Jonah Rockoff.

Consolidated and managed large-scale NYC administrative data to support empirical and theoretical research on the effects of paraprofessionals effectiveness.

Inter-American Development Bank (IDB). **Sep 2023 – Dec 2023**
Consultant – Education Division

Conducted a comprehensive literature review related to the impact of specialized schools. Developed analytical frameworks and methodology recommendations for student selection and classification in public school assignment systems, supporting evidence-grounded project design.

Teachers College, Columbia University. **Sep 2022 – Sep 2023**
Graduate Research Assistant - Professor Sarah Cohodes.

Collaborated on a multi-institutional research project with Harvard, Stanford, and Brown evaluating the long-term impacts of charter schools on educational and labor market outcomes. Managed and cleaned large-scale administrative datasets across multiple state systems.

Inter-American Development Bank (IDB). **Apr 2022 – Sept 2022**
Consultant – Social Sector Gender and Diversity Division

Provided empirical analysis and technical input for the Gender and Diversity Sector Framework Document. Synthesized global evidence on the effectiveness of gender and diversity programs across LAC to inform IDB strategy.

Center for the Economics of Human Development–The University of Chicago. **Nov 2020 – Apr 2022**
Pre-Doctoral Fellow, Professor James Heckman.

Supported the production of empirical and theoretical research guided by Professor James J. Heckman. Contributed to the construction of evidence-based studies related to inequality, development, and investments in early childhood programs that foster skills over the life cycle.

Inter-American Development Bank (IDB). Washington, DC **Sep 2017 – Oct 2020**
Research Analyst –Front Office Social Sector (SCL)

Contributed to the production of analytical studies on teacher quality and professional development interventions, examining medium-term impacts using administrative and field data. Supported design and implementation of a large-scale teacher quality program in Ecuador covering ~28,000 students and 4,000 teachers.

Concentra- Intelligence in Energy. Bogota, Colombia. **Jul 2016 – Aug 2017**
Economic Analyst

Consolidated national datasets and developed demand forecast models. Built indicators related to energy sector and developed reports.

Department of Economics, University of the Andes. Bogota, Colombia. **Nov 2014 –Jul 2016**
Research Analyst **-** Prof Fabio Sánchez

Developed predictive models of university dropout using large-scale administrative data (SPADIES system); produced policy briefs for targeted retention strategies.
Contributed to a World Bank-affiliated project using spatial econometrics and hedonic pricing for residential water valuation.
Supported education and social development research through data analysis, literature reviews, and econometric estimation.

TEACHING EXPERIENCE

- Teaching Assistant, Data Analysis for Policy (Graduate) — Teachers College, Columbia University (2024)
- Teaching Assistant, Economics and Education (Graduate) — Teachers College, Columbia University (2023)
- Teaching assistant, Theory of the Consumer and the Firm (Undergraduate) — Universidad de los Andes (2015-2016)
- Teaching assistant, Microeconomics II (Undergraduate) — Universidad de los Andes (2015 – 2016)
- Head of the econometrics area in the study hall ECOMAMI (Undergraduate) — Universidad de los Andes(2014)

HONORS AND AWARDS

- American Association of University Women (AAUW). International Fellowship -25K. 2026-2027
- Vice President's Grant for Student Research in Diversity - 3K. Columbia University. 2025
- Center for Development Economics and Policy Student Research Grant Award – SIPA, Columbia University. 2024-2025
- Student research grant, Economics and Education Ph.D. Students Award- 1.5 K Columbia University. 2023
- Teachers College Doctoral Research Fellowship. Columbia University. New York. 2022
- Nominated to Juan Luis Londoño Award, Economics Department, Universidad de los Andes. Colombia. 2017

SKILLS

- **Languages**: Spanish (Native), English (Advanced), French (Intermediate), Portuguese (Beginner)
- **Methods**: Causal inference (IV, RD, DID), IRT/psychometrics, impact evaluation, applied micro econometrics, quasi-experimental design, predictive modeling, instructional design, evaluation framework design, generative AI evaluation.
- **Software:** Stata, R, Python, Matlab, ArcGIS, GeoDa, E-Views, MS Office

TEACHERS COLLEGE COLUMBIA UNIVERSITY

TEACHERS COLLEGE
COLUMBIA UNIVERSITY

Student No: T00049248 Date Issued: 07-JUL-2026

Record of: Cindy Alejandra Campos Quintero Page: 1
Issued To: CINDY CAMPOS
AAC2271@TC.COLUMBIA.EDU
DocumentID: TRAKCUIL-80427928

Level: Graduate
Candidacy: Doctor of Philosophy
Dept: Educ Policy & Social Analysis
Current Major: Economics and Education

SUBJ	NO.	COURSE TITLE	CRED	GRD
TRANSFER CREDIT ACCEPTED BY THE INSTITUTION:				
2015-2016		Los Andes Univ COLUMBIA		
TRAN	9001	ADVANCED ECONOMETRICS	4.0	T
TRAN	9002	THESIS SEMINAR	4.0	T
TRAN	9003	INTRO TO SPATIAL DATA ANALYSIS	4.0	T
TRAN	9004	ADV MICROECONOMICS GAME THEOR	4.0	T
TRAN	9005	ADVANCED SHORT-TERM MACROECONO	4.0	T
TRAN	9006	ECONOMIC THEORY OF CONFLICT	4.0	T
TRAN	9007	CONTRACT THEORY	2.0	T
TRAN	9008	METHODS IMPACT EVAL INTERVENTI	4.0	T
		Total Earned Credits 4.00		
Summer Term 2022				
IN PROGRESS WORK				
IND	5000	Independent Research	0.0	
Autumn Term 2022				
G	5073	CU:Machine Learning Soc Sci	3.0	A+
G	5760	CU:Exper Studies For Pols &	4.0	A
BBSN	5007	Neuroscience Apps to Education	3.0	A-
HUDM	4125	Statistical Inference	3.0	B-
Spring Term 2023				
G	6306	CU:Industrial Development	3.0	A-
G	6451	CU:The Economics Of Labor I	3.0	A
EDPA	6002	Quant Methods Eval Ed Polc	3.0	A
EDPE	6050	Educ & Eco Devel:Adv Topics	3.0	A
Autumn Term 2023				
G	6466	CU:Economics Of Education	3.0	A
EDPE	4155	Eval Ed Privatiztn Schl Choice	3.0	A
EDPE	5550	Workshop in Economics in Educ	1.0	YC

********** CONTINUED ON NEXT COLUMN **********

SUBJ	NO.	COURSE TITLE	CRED	GRD
Continued:				
GE	2012	Causal Inference	3.0	B
Spring Term 2024				
G	6307	CU:Pub Economics And Develop	3.0	A
G	6491	CU:Experimental Economics	3.0	A
GE	2524	Psychological Measurement	3.0	A-
IN PROGRESS WORK				
EDPE	5550	Workshop in Economics in Educ	2.0	
Autumn Term 2024				
G	6493	CU:Behavioral Economics	3.0	A
EDP	7500	Dissertation Dsgn & Prep	0.0	P
Spring Term 2025				
EDP	7500	Dissertation Dsgn & Prep	0.0	P
Summer Term 2025				
IN PROGRESS WORK				
EDP	7500	Dissertation Dsgn & Prep	3.0	
Autumn Term 2025				
IN PROGRESS WORK				
EDP	7500	Dissertation Dsgn & Prep	0.0	
Spring Term 2026				
IN PROGRESS WORK				
EDP	7500	Dissertation Dsgn & Prep	0.0	

********** END OF TRANSCRIPT **********

NOT VALID WITHOUT REGISTRAR'S SIGNATURE IN THIS AREA

Christopher Nieves
Registrar

TEACHERS COLLEGE COLUMBIA UNIVERSITY

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
525 W. 120th St., New York, NY 10027

OFFICE OF THE REGISTRAR

This transcript covers coursework taken from Summer 1989 to the present. If prior coursework exists it will appear on a separate, pre-computerized sheet with its own definitions. Note that this transcript is considered official for the listed recipient only.

Non-credit offerings from Spring 2015 forward may also appear. These offerings may be courses, professional development experiences, or instructional programs which do not require extensive assessment or examinations and do not offer academic credit. Students who have taken both non-credit and for-credit courses will receive two separate transcripts for each type of work.

BY AGREEMENT, Teachers College degrees are conferred by Columbia University. Degrees are conferred in October, February, and May. Furthermore, by affiliation with Columbia University, students are permitted to enroll in courses offered by the University. By agreement with Union Theological Seminary and the Jewish Theological Seminary of America, students are permitted to enroll in courses offered by those institutions. The College is authorized by the New York State Education Department to offer programs leading toward the B.S., M.A., M.S., Ed.M., and Ed.D. degrees. In accordance with an Agreement between Columbia University's Graduate School of Arts and Sciences, Teachers College offers the Ph.D. degree in Psychology and in Education.

TEACHERS COLLEGE IS ACCREDITED BY: New York State, Middle States Association of Colleges and Schools, American Psychological Association (Counseling, Clinical, School), American Board of Examiners in Speech Pathology and Audiology of the American Speech and Hearing Association, and Council for the Accreditation of Educator Preparation (CAEP). Individual programs may also be accredited by other professional bodies. Please visit http://www.tc.columbia.edu/catalog/ for more information.

CREDIT: A point is equivalent to a semester hour and is used for graduate-credit courses. Courses offered during a non-credit term do not receive graduate credit. Any numeric units listed next to a non-credit course are considered CEU's. See below.

CONTINUING EDUCATION UNIT: Certain non-credit offerings at Teachers College lead to the awarding of continuing education units or CEU's. CEU's awarded, if any, will be listed next to each eligible course on the transcript. Teachers College is authorized by the International Association of Continuing Education and Training (IACET) to award continuing education units (CEU's). One CEU equals ten contact hours of attendance and participation in organized continuing education and/or training experiences under responsible, qualified direction and instruction.

GRADES *(effective September 1984)*: A+, Rare performance, highly exceptional; A, Excellent, outstanding; A-, Excellent work but not quite outstanding; B+, Very good, solid achievement expected of most graduate students; B, good, acceptable achievement; B-, Acceptable achievement, but below what is generally expected of graduate students; C+, Fair achievement, above minimally accepted level; C, Fair achievement, but only minimally acceptable; C-, Very low performance; F, Failure; P, Passed; DP, Doctoral Pass; WD, withdrawn; YC (Year course) indicating first half, upon completion of second half a grade is recorded validating both terms of registration; R, Attendance credit; IN, Incomplete; *, Grade not reported.

Students in non-credit courses will receive grades of either satisfactory (S) or unsatisfactory (U). Some non-credit courses or programs may be graded using traditional letter grades. If a grade of S or a passing letter grade is issued, the College assumes that the student has attended all meeting hours of a given non-credit offering. If the student receives a grade of S, the student will be eligible for all CEU's associated with the course (provided the course is approved for CEU's). A grade of "*" means that a grade has not been reported by the course instructor.

GRADE POINT AVERAGE AND CLASS RANK: Teachers College does not calculate a grade point average, rank its students, or award honors at graduation. **Unless otherwise noted, a student/graduate is considered to be in good standing.**

PRIVACY ACT: This transcript has been forwarded to you with the understanding that it will not be released to any other parties, in conformity with the Family Educational Rights and Privacy Act of 1974.

COURSE NUMBERING: *Effective Autumn 1998*, the first three letters of a course prefix identify Department; all four letters together indicate subject area. First numeral indicates course credit level. Teachers College affiliates, including Columbia University, Barnard College, Union Theological Seminary, and Jewish Theological Seminary, as well as member schools of the Interuniversity Doctoral Consortium, do not follow this numbering convention:

Department Codes: A&H, Arts & Humanities; BBS, Biobehavioral Sciences; CCP, Counseling & Clinical Psychology; C&T, Curriculum & Teaching; EDP, Education Policy & Social Analysis; HBS, Health & Behavioral Studies; HUD, Human Development; ITS, International & Transcultural Studies; MST, Mathematics, Science & Technology; ORL, Organization & Leadership; IND, Interdepartmental.

Course Credit Level: 4 initial graduate; 5 intermediate graduate; 6 advanced graduate; 7 doctoral dissertation seminar; 8 doctoral dissertation advisement; 9 post-doctoral study.

Non-credit courses will use the same department and subject area codes but will not follow the same numbering conventions as credit courses and will appear in distinct, non-credit academic terms and on separate transcripts.

Hidden Strengths: Adversity, Adaptation, and the Measurement of Skills*

Alejandra Campos Quintero†

Working Paper - Do not circulate

September, 2026

Abstract

Research in psychology, evolutionary anthropology, and, more recently, economics suggests that individuals raised in harsh and unpredictable environments develop stress-adaptive skills that help them cope with adversity. This study examines whether conventional cognitive assessments underestimate the abilities of students from high-adversity backgrounds, by overlooking the contexts in which their skills are most likely to manifest. Specifically, I examine whether exposure to environmental noise during testing—a real-world stressor—differentially affects student performance depending on childhood adversity. To test this, I combine individual-level data from Colombia's high school exit exam (SABER 11) with georeferenced noise pollution data from the capital city. Leveraging the quasi-random assignment of students to test centers, I find that students from high-adversity backgrounds perform relatively better in noisy settings, suggesting greater resilience to environmental stressors. The results highlight the importance of considering testing conditions in educational evaluation, and raise concerns about the potential allocative inefficiencies of having high reliance on noisy signals of student quality.

Keywords: Stress-adaptive skills; Human Capital; Skills; Adversity; Noise pollution

*I am grateful for valuable discussions and comments from Alex Eble, Carolina Concha-Arriagada, Jonah Rockoff, Sebastian Otero, Fabio Sanchez, Mark Dean, Michael Best, Alessandra Casella, Lisa Ho, Ofer Malamud, Alejandro Ganimian, Judd Kessler, Donald Green, Saloni Gupta, Elizabeth Jaramillo, and Alejandra Quintana.

†Columbia University. aac2271@tc.columbia.edu.

I Introduction

It is widely acknowledged that early exposure to disadvantaged environments has significant detrimental effects on the developmental trajectories of children (Almond and Currie, 2011; Doyle, 2020; Mogstad and Torsvik, 2023). While much of the economics literature has traditionally focused on understanding and mitigating the negative effects of childhood adversity, there is a growing body of research in psychology, anthropology, and, more recently, in economics exploring potential positive adaptive responses to these environments.[1]This emerging research suggests that individuals raised in harsh and unpredictable environments can develop a specialization of skills optimized to cope with high-adversity contexts (Frankenhuis and de Weerth, 2013; Ellis and Del Giudice, 2019; Wantchekon and Zhang, 2024). However, these skills, while valuable in the real-world and labor market contexts (Deming, 2017; Jackson, 2018), may only become apparent under specific conditions (Mittal et al., 2015).[2]

This paper investigates the context-sensitive expression of ability, and how the environment of evaluation influences the signals of student ability. Specifically, I ask: Does exposure to environmental stressors during high-stakes testing differentially affect performance based on a student's childhood adversity? I address this by leveraging a quasi-natural experiment in Colombia, combining individual-level data from the national high school exit exam (SABER 11) with georeferenced noise pollution data in Bogotá. Using the quasi-random assignment of students to test centers, I identify the interaction between ambient noise—a real-world stressor—and childhood adversity. This empirical analysis is complemented by a companion laboratory experiment designed to isolate the non-cognitive mechanisms and test the malleability of these adaptive responses through a targeted intervention.[3]

Building on this framework, my empirical analysis produces two primary findings. First, I demonstrate that while environmental noise generally impairs test performance, students from high-adversity backgrounds exhibit a relative resilience absent in their more advantaged peers. Specifically, being from a high-adversity background—defined as the bottom quartile of the Living Con-

[1]Primarily analyzing the impact of early childhood interventions related to health care, pre-school programs, and targeting home environments (Duncan et al., 2023). One exception is Wantchekon and Zhang (2024), which provides evidence that adverse economic conditions, such as weather shocks, promote the development of noncognitive skills like self-efficacy, self-esteem, and an internal locus of control.

[2]Among some of the stress-adaptive skills identified in this literature are attention shifting (Blair and Raver, 2014; Evans et al., 2013; Mittal et al., 2015), working memory updating (Young et al., 2018), effective group process (Dittmann et al., 2020), creativity (Dahlman et al., 2013; Fry, 2018), self-efficacy, self-esteem, and an internal locus of control (Wantchekon and Zhang, 2024).

[3]This draft presents only results from the empirical analysis using administrative data. In the laboratory experiment, designed to test mechanisms in a more structured setting, I ask: Do stress-adaptive skills such as attention shifting, working memory updating, creativity, resilience, and adaptive intelligence help them cope under these conditions? Can adaptability itself be strengthened through short interventions? Findings from the laboratory experiment will be incorporated in subsequent versions of the paper.See Appendix IX. E.

ditions Index—is associated with a 0.025 standard deviation relative improvement in performance when testing in noisy versus quiet environments. Second, I show that these environmental conditions have high-stakes downstream consequences: using a quasi-experimental approach, I find that assignment to a higher-quality test center significantly increases the probability of both college enrollment and graduation. Together, these results suggest that standardized testing environments may systematically mismeasure the latent ability of disadvantaged students.

Causally identifying whether performance gaps across socioeconomic groups reflect differences in underlying skills or the context in which skills are expressed presents several empirical challenges. First, socioeconomic status is highly correlated with parental investment, income, and school quality, all of which influence academic performance. Consequently, comparisons between groups cannot disentangle latent skill gaps from environmental differences. Second, testing conditions are rarely assigned at random, and while standardized exams are designed to minimize variation across sites, students from lower-SES backgrounds are disproportionately exposed to adverse testing environments, limiting the scope for identifying how performance responds to environmental stress.

I overcome these challenges by exploiting a unique feature of the SABER 11 registration process. During registration, students indicate a preferred testing zone from among ten designated areas in the city. While some students are placed within their preferred zone, capacity constraints force a subset of students to be reassigned to available seats in other zones.[4] For this reassigned group, the specific level of noise pollution at the assigned test center is plausibly exogenous, conditional on their initial zone of preference. I estimate intent-to-treat (ITT) effects by regressing standardized SABER 11 scores on test-center noise exposure, including preferred-zone fixed effects and a rich set of predetermined student covariates. To account for within-site correlation, I cluster standard errors at the testing-center level.

Consistent with the identifying assumption, Table 1 shows that student characteristics are well-balanced across noise levels within this reassigned sample. Across nearly all dimensions—including gender, household composition, and maternal education—there are no statistically significant differences between students assigned to noisy versus quiet test centers. While the large sample size generates a minor statistical difference in age, the magnitude is negligible (less than 0.05 years), and I find no evidence of systematic selection into testing environments based on socioeconomic status. These results confirm that, conditional on the zone of preference, noise exposure is orthogonal to student background.

[4] After priority placements (e.g., for students with disabilities or ethnic groups), students are assigned to a center within their preferred zone if capacity allows; otherwise, they are reassigned to centers with available seats in other zones.

I proxy for childhood adversity using a multidimensional living conditions index constructed from the SABER 11 survey. This index summarizes household access to basic resources, parental characteristics, and household assets, serving as a consistent measure of socioeconomic disadvantage. Students in the bottom quartile of this index are classified as coming from a high-adversity background. These students typically face multiple constraints, including limited access to technology and learning materials, lower parental education, fewer household assets, and, in many cases they need to contribute to household income through work.

Additionally, environmental noise is measured using predictions from an acoustic propagation model developed by the Environmental Secretariat of Bogotá. The model is calibrated with data from 36 fixed monitoring stations across the city during 2018–2019 and estimates how sound travels through and interacts with the urban environment by integrating information on geography, infrastructure, traffic flows, and environmental conditions[5]. Importantly, the model provides standardized estimates of noise levels for Sundays between 7 a.m. and 1 p.m.—the time window during which SABER 11 exams are administered. Although it does not capture real-time fluctuations, the modeled approach reduces measurement error relative to short-distance recordings and allows me to have coverage of all test centers. Each center is assigned the estimated decibel level of the grid cell in which it is located.[6]

Using these measures, I provide empirical evidence that students from high-adversity backgrounds exhibit are less disturbed by noisy testing environments, consistent with the hypothesis of stress-adaptive skills (see Figure 4). Specifically, I find that assignment to a test center with noise levels above 65 dB is associated with a modest increase of 0.02 standard deviations in performance for students from a background with a higher level of adversity[7]. While small in magnitude, this effect is comparable to estimates in the literature on environmental influences such as air pollution and heat (Park et al., 2020; Kremer et al., 2024). These results suggest that environmental conditions can shape the expression of cognitive skills, with important implications for the validity of standardized test scores.

To further explore heterogeneity in this effect, I disaggregate the analysis by quartiles of adversity. The results suggest that the relative gains from noise exposure among students from more disadvantaged backgrounds is nonlinear. Estimated effects increase across quartiles and are the largest in the third, while remaining positive but somewhat smaller in the highest quartile. This

[5]The outputs are generated at a resolution of 10×10 meters, ensuring detailed spatial coverage.

[6]Bogotá's Strategic Noise Maps (MER) are developed in accordance with international best practices for environmental noise modeling, drawing on methodologies consistent with the European Union's Environmental Noise Directive (2002/49/EC), which sets standards for noise mapping and urban acoustic assessment.

[7]World Health Organization (1999)'s recommended limit inside the classroom is 35 dB. Exposure to levels above 55-60 dB can be considered disturbing.

inverted U-shaped pattern is consistent with models in which moderate adversity fosters adaptive responses to environmental stressors, but very high adversity eventually impairs cognition. This pattern highlights the dual role of adversity as both a potential source of adaptability and a constraint when exposure becomes extreme. To rationalize these results, I present a simple theoretical framework in section VII.

As a complementary exercise, I also explore alternative proxies for adversity: (i) living in a crowded household (more than six members), (ii) living in a single-parent household, and (iii) working at the time of the exam. Results using the first two measures—crowding and student work—are consistent with and slightly larger than those obtained with the composite adversity index, suggesting that direct exposure to competing demands and unpredictability may strengthen adaptive responses. In contrast, there is no evidence of a differential effect of noise for students from single-parent households, indicating that not all dimensions of adversity operate through the same mechanisms. Taken together, these results provide new evidence on how testing environments interact with students' backgrounds to shape cognitive performance.

To assess robustness, I conduct several complementary checks. A primary concern is that the main effect decreases once school fixed effects are included. Under truly random assignment within schools, including fixed effects, should increase precision, rather than drive the coefficient toward zero. To assess this, I simulate the assignment process under random allocation and show that the observed attenuation is consistent with the limited within-school variation in treatment exposure, rather than a violation of the identification assumption. I further restrict the sample to schools with sufficient variation in both adversity and noise exposure; the results remain broadly similar, with the estimated noise effect turning negative and significant. Finally, the findings are robust to alternative clustering levels and to restricting the sample to test centers located near noise monitoring stations. Altogether, the evidence suggests that although standardized tests play an important role in educational evaluation, they may also provide a noisy signal of underlying student potential.

While this quasi-experimental section provides valuable insights into how external conditions influence cognitive performance, it cannot fully disentangle whether noise impairs performance among low-adversity students, activates resilience among high-adversity students, or both.[8] To address this limitation, the second component of the study—a controlled lab experiment—will provide direct causal evidence on whether individuals from high-adversity backgrounds demonstrate stronger or weaker performance across various skills when exposed to adversity—and whether this

[8]For instance, noisy environments may reduce the performance gap between students from low- and high-adversity backgrounds by disproportionately affecting low-adversity students, triggering adaptive skills in high-adversity individuals, or through a combination of these mechanisms.

differs from their performance under standard conditions.[9] Together, the two components form a complementary strategy: the quasi-experimental analysis establishes the real-world relevance of context-sensitive performance differences, while the lab experiment allows me to test the mechanisms underlying these differences in a controlled environment. This paper focuses specifically on the quasi-experimental component to motivate the lab experiment.

Overall, the findings suggest that standardized tests, by not accounting for context-sensitive abilities, may provide systematically biased signals of student potential. Such mismeasurement limits our understanding of human capital formation and can reinforce socioeconomic inequalities by misallocating educational and occupational opportunities (Ebenstein et al., 2016). By combining quasi-experimental evidence from a high-stakes testing environment with a planned laboratory experiment designed to isolate mechanisms, the paper aims to advance our understanding of how adversity shapes skill development and expression. More broadly, the results underscore the need for evaluation and pedagogical frameworks that recognize the adaptive dimension of skills and that better reflect the diverse environments in which students grow up.

This paper is organized as follows. Section II summarizes the current literature and contributions. Section III will describe the context and data. Section IV will describe the empirical strategy to evaluate whether the environment matters, using administrative data from Colombia. Section VI and VII shows the results and robustness checks, respectively. Section VIII provides a simple theoretical framework for the development and elicitation of stress-adaptive skills. Finally, Section IX, concludes.

II Related literature and Contributions

A long-standing literature in psychology and anthropology documents that early exposure to harsh or unpredictable environments can foster specialized "stress-adapted" skills. These include heightened empathic accuracy (Kraus et al., 2012), sensitivity to social cues (Brienza and Grossmann, 2017), and improved performance in collaborative or collective tasks (Dittmann et al., 2020). While adversity is often framed through a deficit model, emerging research indicates that individuals raised in high-stress contexts may develop a comparative advantage in flexible task-

[9] The experiment will assess both standard cognitive abilities and stress-adaptive skills—attention shifting, working memory updating, creativity, resilience, and adaptive intelligence—by randomly assigning participants from diverse childhood backgrounds to different testing environments. Randomization will be stratified based on demographic factors, including socioeconomic status and childhood adversity, as measured by the Adverse Childhood Experience (ACE) Questionnaire for Adults. Each participant will be assessed under both a structured and an adverse environment. In the structured condition, tasks are completed in a quiet, predictable setting with stable rules and orderly sequencing. In the adverse condition, participants face calibrated disruptions— background noise, visual clutter, screen lags, and unpredictable time reminders—designed to mimic real-world instability. The order of exposure to these environments is individually randomized, allowing for clean identification of whether and how performance varies across contexts within the same individual. The details of this experiment, however, are beyond the scope of this paper and will be presented separately.

switching and tracking novel information when evaluated under environmental pressure(Young et al., 2018).[10]

My work contributes to several strands of literature. First, the traditional human capital accumulation literature, which finds that adversity negatively affects skill accumulation unless mitigated by parental overcompensation (Cunha and Heckman, 2007; Heckman and Kautz, 2013) or shifts in opportunity costs of education (Bau et al., 2020). I extend this by proposing that adversity does not merely diminish skills but shifts the specialization of the skill toward traits optimized for high-stress contexts. This aligns with recent causal evidence showing that scarcity can actually improve certain types of decision-making by increasing the utility loss of forgone opportunities, and adverse shocks can promote non-cognitive traits like self-efficacy (Fehr et al., 2022; Wantchekon and Zhang, 2024). By documenting that high-adversity students exhibit relative resilience under environmental stress, I provide evidence of a relationship between environment and skill expression currently absent from standard human capital models (see Appendix VII for a theoretical framework).

Second, this paper connects to the Fetal Origins Hypothesis and the literature on the biological embedding of early-life shocks (Almond and Currie, 2011). While much of the health economics literature focuses on the effects of adversity—where early-life deprivation leads to persistent deficits in cognitive and physical development—I build on an emerging perspective that views these responses as phenotypic adaptations. This approach, recently evidenced in the context of interrupted malaria control in Zambia (Fink et al., 2021), suggests that biological systems may optimize human capital investments to suit high-adversity environments rather than simply suffering structural damage. I contribute to this field by showing that these biological adaptations have measurable economic consequences. Specifically, my findings suggest that when the evaluative environment "matches" the student's developmental context (e.g., a high-stress, noisy test center), it can reveal specialized cognitive advantages that remain latent in standard settings.

Third, I contribute to the debate regarding the validity of standardized tests and measurement bias. Research documents significant performance gaps across socioeconomic backgrounds (Banerjee et al., 2025). Additionally, it has been found that specific testing conditions—such as the competitive nature of the exam or the format—can differentially penalize groups (Baldiga, 2014). I show that the physical environment of evaluation is a critical source of measurement bias. While Jihao Hong Justin (2025) find that testing in a different school than your own disproportionately hurts low-SES students, I show that within those environments, certain stressors like noise can reveal latent abilities. By investigating whether adaptability functions as a mediator that enables students

[10]These findings align with evolutionary perspectives that link the development of cognitive abilities to ecological challenges (Rosati et al., 2023)

to maintain performance across contexts, I extend resilience research toward the fundamental problem of signal noise in standardized assessments.

Additionally, this project connects to a growing literature on how poverty-related stimuli—such as ambient heat, air pollution, hunger, sleeplessness, and financial strain—affect productivity (Graff Zivin and Neidell, 2012; Graff Zivin et al., 2020; Chang et al., 2019, 2016; Park et al., 2020; Adhvaryu et al., 2020; Schilbach, 2019; Bessone et al., 2021; Dean, 2024). I extend this work by examining heterogeneity in these effects by adversity background, testing whether individuals with higher predicted levels of stress-adaptive skills are differentially affected by environmental shocks in performance tasks.This suggests that the optimal design of evaluation environments may need to account for the developmental history of the individuals being assessed.

Finally, my dissertation contributes to the literature on the value of socio-emotional skills and allocation of talent in the labor market. Prior work shows that traits such as teamwork and creativity are increasingly rewarded (Deming, 2017; Jackson, 2018). However, the lack of information on the skills of workers at recruitment affects matching (Bassi and Nansamba, 2021). I build on this literature by linking the mismeasurement of stress-adaptive skills to misallocation in education. Specifically, I highlight how noisy signals of ability can distort sorting, overlook latent abilities, and can affect allocation into college.

III Context and Data

Standardized assessments are a central tool of current education systems and labor markets, providing benchmarks for academic progression and workforce placement. However, when these assessments fail to capture the full spectrum of human abilities, they risk distorting the measurement of human capital, misallocating opportunities, and reinforcing barriers to social mobility (Ebenstein et al., 2016). One setting where this concern is especially relevant is among students from high-adversity backgrounds, where conventional cognitive assessments may systematically mismeasure ability by failing to account for the contexts in which stress-adaptive skills are most likely to emerge.

This project examines this concern in Bogota, Colombia, where high-stakes standardized exams determine access to higher education and scholarships. Using administrative data for 2018–2019 from the Colombian Ministry of Education, the Colombian Institute for the Evaluation of Education (ICFES), and the Environmental Observatory of Bogota, I draw on a comprehensive dataset covering 107,734 students who sat for the national high school exit exam across nearly 400 test centers. The statistics for my sample are presented in Table 1 in two panels. Panel A reports sample means of my student-level data, and Panel B reports sample means at the school-level data.

III. A Standardized testing data and test center

The Colombian Institute for the Promotion of Higher Education (ICFES)[11] administers the national high school exit exam (*SABER 11*), a high-stakes standardized test taken by more than 90% of high school seniors in Colombia, regardless of their post-secondary plans. The exam evaluates students across several subjects, including language, mathematics, social studies, and the natural sciences. *SABER 11* scores are used both for college admissions and for assessing the quality of secondary education nationwide. I use data from ICFES to measure academic achievement, identify students' test center locations, and I leverage the random assignment of students to test centers as part of my identification strategy[12]. In particular, I will use individual level data from the four test administrations between 2018 to 2019. The dataset also includes demographic variables such as gender, age, and employment status at the time of testing. To ensure comparability across test administrations, I normalize student scores by subject and cohort.

III. B Adverse upbringing

Childhood adversity refers to a broad range of negative experiences in childhood, which may include food insecurity, child maltreatment (i.e., abuse, neglect), parental incarceration, and exposure to intimate partner violence. McLaughlin et al. (2014) developed the dimensional model of adversity and psychopathology (DMAP), which proposes two dimensions of childhood adversity: experiences that deprive children of environmental stimulation and resources to promote child development (Deprivation) and experiences that threaten the child's physical and emotional safety (Threat).

Each of these forms play a distinct role in shaping individual experiences and developmental pathways. For this study, I will focus on two indicators of deprivation, using proxies to measure their impact: economic hardship, household structure[13].

i. Economic adversity

Economic adversity, characterized by poverty and limited access to basic needs, poses significant barriers to children's development. Children that grow up in disadvantaged households often face constraints that hinder the acquisition of both cognitive and non-cognitive skills, leading to persistent educational and socioeconomic disadvantages over time (Dean et al., 2017).

To proxy for economic adversity, I use the Living Conditions and Insecurity Index (LCI), developed by ICFES. The LCI is based on detailed information about household characteristics, including access to basic services, housing quality, and the ownership of durable goods. I define a binary

[11] *Instituto Colombiano para la Evaluación de la Educación*

[12] See Section IV for details on the randomization procedure.

[13] I will also include a measure of violence and conflict.

indicator of adversity that equals one if a student falls within the bottom 25th percentile of the LCI distribution (See Figure .

ii. Household structure

Household structure is a key dimension of adversity. Children raised in single-parent or overcrowded households often face distinct challenges, including limited access to supportive networks and increased caregiving burdens. Using data from the ICFES demographic survey, I construct two indicators of household structure: a binary variable identifying students from single-parent households, and another indicating whether the student lives in a crowded household (defined as having more than six people).[14]

III. C Noise data: proxy to elicit unpredictable environment

Bogotá's Environment Secretariat (ES) established the Environmental Noise Monitoring Network (*Red de Monitoreo de Ruido Ambiental de Bogotá*, RMRAB), a real-time acoustic monitoring system composed of 36 stations distributed across the city. The network operates in 18 of Bogotá's 20 localities. It is the largest environmental noise monitoring system in Colombia and one of the most important in Latin America, designed to continuously track and assess urban noise conditions across the Capital District.

Figure 1 shows the variation in noise pollution levels across Bogotá's city grid, based on hourly daytime noise data from Sundays between 2018 and 2019. The estimates of the Environmental Noise Map (MER) are produced by the Environment Secretariat (ES), using an acoustic propagation model that simulates how sound travels through and interacts with the urban environment. The model integrates information on urban geography, noise sources, and environmental conditions to estimate spatial patterns of noise exposure.[15]

I assign each test center the corresponding noise level from the grid cell in which it is located, which are represented as blue dots in Figure 1. Using this spatially linked dataset, I construct my main variable of interest: a binary indicator for ***noise exposure***, which takes the value of 1 if the test center is situated in an area with average noise levels above 65 decibels (dB).[16] For each test administration between 2018- 2019, I also know the date of the test and the precise location of the test center where the exam is administered, allowing me to also use the raw noise pollution

[14] As a next step, I plan to complement this information with household-level data from SISBEN (the Identification System for Subsidies Beneficiaries), a nationwide means-testing registry established in 1994. SISBEN data includes information on demographics, housing, income, and employment. It will allow me to identify household locations and characteristics of all household members.

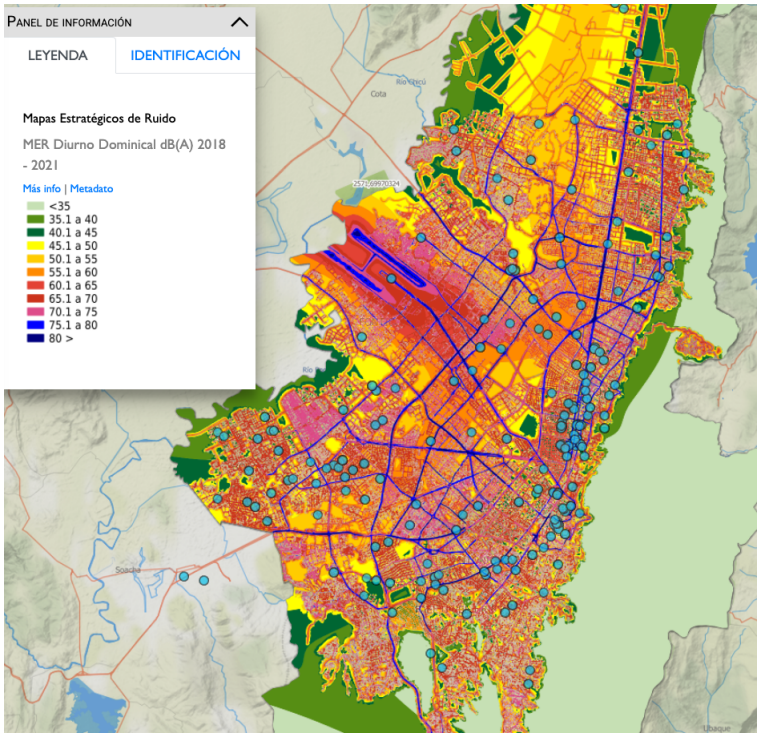
[15] See Appendix Section IX. C for further details.

[16] World Health Organization (1999)'s recommended limit inside the classroom is 35 dB. Exposure to levels above 55-60 dB can be considered disturbing.

measure to each test administration[17]. However, I will use as a preferred measure the estimations from the noise propagation model provided by the ES.

Figure 2 shows the mean noise levels and the mean Living Conditions Index (LCI) for each test center, disaggregated by geographic zone. Each green dot represents the ambient noise level (dB) at a specific center, while the purple diamonds indicate the average LCI of the students assigned to that center. The figure reveals two patterns: first, it shows substantial intra-zone variation in noise exposure, providing the necessary "treatment" variance to identify the effects of environmental stress. Second, and most importantly, it shows that there is no systematic correlation between a center's noise level and the average socioeconomic background of its students within a given zone. For instance, in zones like the North or Southwest, centers with vastly different noise profiles (ranging from approximately 55 to over 70 dB) serve student populations with nearly identical mean LCI scores. This lack of clustering suggests that the assignment mechanism successfully decoupled environmental testing conditions from student adversity, supporting the claim that noise exposure is plausibly exogenous conditional on the zone of preference.

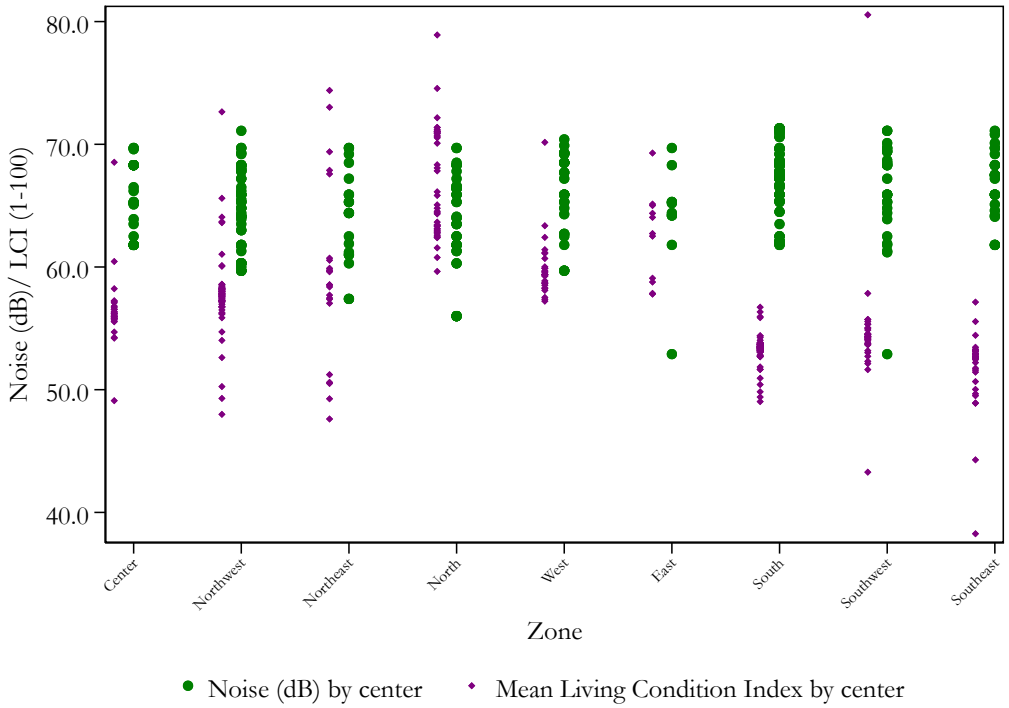
Figure 1. Noise pollution and test centers Bogota, Colombia



Notes: The blue dots correspond to a test center location from 2018-2019. *Source*: ES and ICFES (2024)

[17]This is a robustness test that I plan to implement.

Figure 2. Mean Noise (dB) and Living Conditions Index (LCI) by test center



Notes: Each dot represents the mean LCI or Noise level of a test center for all periods (2018-2019). Lower levels of LCI represent more adversity. Figures 13 and 14 disaggregate it by period.

IV Empirical strategy

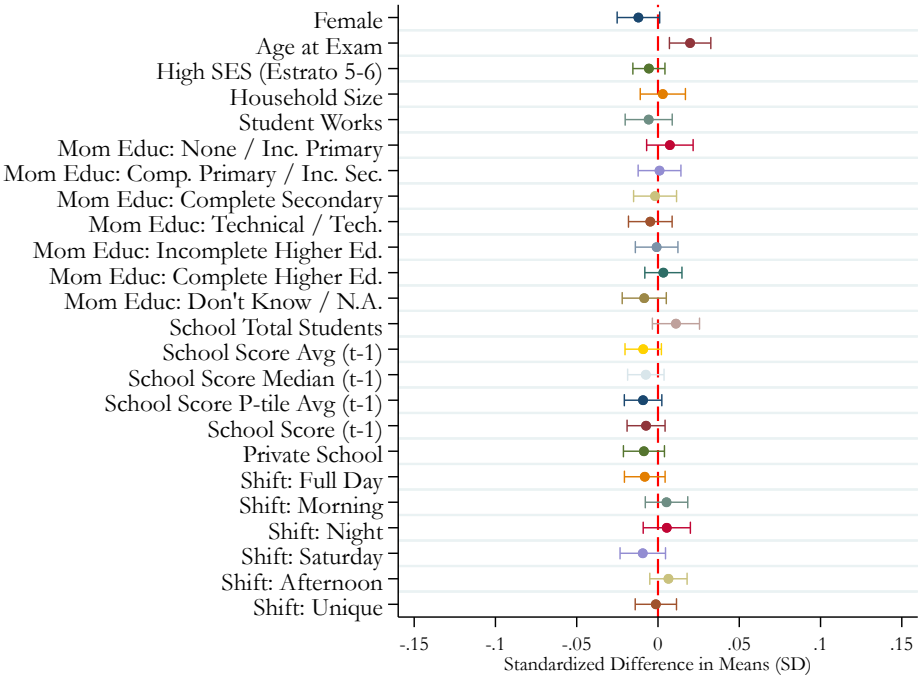
This study aims to integrate a quasi-experiment and a lab experiment to examine the role of environmental stress-adaptive skills in educational and cognitive performance. The empirical component, although more limited, provides external validity by examining how testing conditions influence cognitive performance in real-world settings, while the lab experiment delves deeper into (1) which specific cognitive and socio-emotional skills are enhanced by exposure to adversity? (2) Does the expression of these skills depend on specific environmental conditions? While both approaches are essential for addressing the research questions, this paper will focus on the first part of this study.

IV. A Empirical strategy: Quasi-experimental Approach

I use administrative data from Colombia's standardized high school exit exam, *Saber 11*, to explore how contextual factors, such as noise pollution, influence skill assessments. The empirical strategy leverages the quasi-random assignment of students to test locations, combined with estimated noise pollution levels in Bogotá at the time of testing.

Students register for the *SABER 11* national high school exit exam and indicate their preferred testing zone from among ten designated areas in the city. Assignment to a specific test center occurs in two steps (Figure 10). First, students with documented disabilities or from an ethnic group are manually assigned to a center within their preferred zone to ensure accessibility. For all other students, if there are available seats in their selected zone, they are randomly assigned to one of the remaining test centers within that area. If no seats remain, students are randomly assigned to a center elsewhere in the city. Because this process involves random assignment conditional on zone capacity, it generates quasi-random variation in exposure to environmental conditions, such as noise pollution.

Figure 3. Probability of being assigned to noisy test center by Treatment Status



Notes: This figure displays the standardized differences in means for baseline student and school characteristics between students assigned to noisy and quiet test centers. The sample is restricted to the students who were assigned to a test center outside of their preferred zone due to capacity constraints. Point estimates are obtained from separate regressions of each standardized covariate on an indicator for assignment to a noisy test center, conditioning on period-by-preferred-zone fixed effects. Standard errors are clustered at the test center level, with 95% confidence intervals represented by the horizontal bars. An omnibus F-test for the joint orthogonality of all 25 covariates yields an F-statistic of 1.46 ($p = 0.086$), failing to reject the null hypothesis of conditional random assignment at the 5% level.

I leverage this feature of the allocation system by focusing on the subset of students who were assigned outside their zone of preference due to capacity constraints. Conditional on preference

zone, for this sample of students, assignment to a noisy center can be treated as random[18]. To verify that capacity constraints induced quasi-random variation in test center assignment, we restrict our sample to students assigned outside their preferred zone. Figure **??** presents the standardized differences in baseline covariates between students assigned to noisy and quiet centers, conditional on testing period and preferred zone. All standardized differences fall well below the standard 0.10 threshold, indicating excellent balance. Furthermore, an omnibus F-test confirms that these 25 student and school characteristics do not jointly predict assignment to a noisy center at the 5% significance level ($F = 1.46$, $p = 0.086$). This strongly supports the assumption that, conditional on period and preferred zone, overflow assignment is effectively random. Similarly, Table 1 compares the means of key covariates between students in high- and low-noise centers. The variables are well balanced, reinforcing the validity of the identification strategy. This design allows me to estimate the causal effect of noise exposure on student performance using the following model:

$$Y_{islt} = \lambda_0 + \gamma D_{lt} + \lambda_1 Adv_i + \alpha(Adv_i * D_{lt}) + \lambda_2 Z_i + \lambda_3 X_i + \varphi_s + \tau_t + \eta_l + \varepsilon_{islt} \quad (1)$$

Y_{islt} represents the test score of individual i from school s, who presented the test at location l in period t. $D_{l(i)}$ is a binary variable that indicates whether the individual was assigned to a test location (l) with high noise levels (above 65 decibels), Adv_i is a dummy that takes the value of 1 if the child comes from an adverse background (Living Conditions Index below 25th percentile), and X_i includes individual characteristics such as age, gender, and distance from school to the assigned test center. The term Z_t is the reported zone of preference used by the centralized algorithm to randomly allocate students to testing centers. Finally, η_l is a vector of test center characteristics[19], and φ_s represents school fixed effects or school value added in previous years, depending on the specification.[20]

[18] Since the probability of getting assigned outside their zone of preference depends on the zone capacity, I still control for zone of preference.

[19] Currently, this vector only includes whether the test center is a school, a university, or another type of facility. I can include this term because the test centers vary by period.

[20] Value added is estimated using results from SABER 11 from the previous years (2016 and 2017) as in (Neilson, 2021).

Table 1. Summary Statistics: Balance Table

Variable	(1) Control: not noisy env. N	(1) Mean/(SE)	(2) Noisy environment N	(2) Mean/(SE)	(1)-(2) Pairwise t-test N	(1)-(2) Mean difference
A. Demographics						
NSE indicator	47626	57.523	58500	56.254	106126	1.269
		(0.043)		(0.038)		
Adversity Index (NSE$\leq p(25) = 1$)	47491	0.231	58285	0.267	105776	-0.037
	128	(0.009)	159	(0.009)	287	
Female	47570	0.526	56975	0.531	104545	-0.005
		(0.002)		(0.002)		
Age	47427	16.945	57862	16.994	105289	-0.049***
	129	(0.014)	160	(0.015)	289	
Socioeconomic status>4	48459	0.055	59275	0.040	107734	0.015
	130	(0.010)	160	(0.008)	290	
More than 6 people in the HH	47510	0.090	58038	0.101	105548	-0.011
	123	(0.003)	153	(0.003)	276	
Student Works	48459	0.253	59275	0.265	107734	-0.013
	130	(0.005)	160	(0.005)	290	
Mom Education=None / Primary Incomplete	46565	0.088	56262	0.101	102827	-0.012
	123	(0.003)	153	(0.004)	276	
Mom Education=Primary Complete / Secondary Incomplete	46565	0.201	56262	0.223	102827	-0.021
	123	(0.006)	153	(0.006)	276	
Mom Education=Secondary Complete	46565	0.251	56262	0.267	102827	-0.015
	123	(0.006)	153	(0.005)	276	
Mom Education=Technical / Technological	46565	0.159	56262	0.159	102827	-0.000
	123	(0.003)	153	(0.003)	276	
Mom Education=Incomplete Higher Ed	46565	0.040	56262	0.035	102827	0.005
	123	(0.001)	153	(0.001)	276	
Mom Education = Complete Higher Ed	46565	0.238	56262	0.197	102827	0.042
	123	(0.015)	153	(0.013)	276	
Mom Education =Don't Know / Not Applicable	46565	0.022	56262	0.020	102827	0.002
	123	(0.001)	153	(0.001)	276	

Notes. This table reports the mean values and standard errors (in parentheses) for student characteristics across test centers categorized by noise levels. "Noisy" centers are defined as those with ambient noise above 65 dB. The final column reports the difference in means and the significance levels from a pairwise t-test. All comparisons include period-zone fixed effects. Standard errors are clustered at the test-center level. Significance: ***=.01, **=.05, *=.1.

	(1)		(2)		(1)-(2)	
	Control: not noisy env.		Noisy environment		Pairwise t-test	
Variable	N	Mean/(SE)	N	Mean/(SE)	N	Mean difference
B. School Characteristics						
Students per school	45174	132.913	54945	145.368	100119	-12.455
	130	(2.422)	159	(1.816)	289	
Mean Saber 11 Score on (2016)	45174	279.602	54945	274.015	100119	5.586
	130	(1.245)	159	(1.120)	289	
Median Saber 11 Score on (2016)	45174	280.152	54945	274.440	100119	5.712
	130	(1.270)	159	(1.142)	289	
Quantiles of Saber 11 Score mean (2016)	45174	73.429	54945	70.042	100119	3.387
	130	(0.747)	159	(0.743)	289	
Quantiles of Saber 11 Score median (2016)	45174	73.530	54945	70.216	100119	3.314
	130	(0.736)	159	(0.733)	289	
Private School	48459	0.529	59275	0.467	107734	0.062
	130	(0.017)	160	(0.016)	290	
School shift - Complete	48459	0.404	59275	0.340	107734	0.064
	130	(0.016)	160	(0.014)	290	
School shift - Morning	48459	0.272	59275	0.308	107734	-0.036
	130	(0.008)	160	(0.006)	290	
School shift - Night	48459	0.060	59275	0.064	107734	-0.004
	130	(0.003)	160	(0.003)	290	
School shift - Saturday	48459	0.039	59275	0.046	107734	-0.007
	130	(0.003)	160	(0.003)	290	
School shift - Afternoon	48459	0.146	59275	0.156	107734	-0.011
	130	(0.006)	160	(0.006)	290	
School shift - Unique	48459	0.081	59275	0.086	107734	-0.005
	130	(0.005)	160	(0.004)	290	

Notes. This table reports the difference in means between low and high noise test centers. High noise pollution is defined as test-centers above 65 dB of noise. All specifications include period-zone fixed effects. Significance: ***=.01, **=.05, *=.1. Standard errors clustered at the center level.

V Results

The results suggest that, under standard conditions, students from low-adversity backgrounds outperform their high-adversity peers on cognitive tests (0.22 sd). However, in noisy environments, students from high-adversity backgrounds appear to exhibit adaptive skills that narrow this performance gap. Specifically, the interaction term (α) in Equation 1 is positive and statistically significant (0.02 sd), suggesting that environmental uncertainty may trigger adaptive responses among high-adversity students.

Table 2, presents multiple specifications to estimate the effect of being assigned to a noisy test center on test performance. Column (1) presents the baseline correlation between noise and test scores. Column (2) adds controls for zone of preference, and Column (3) clusters standard errors at the test-center level. Column (4) introduces the key interaction between noise exposure and adversity. Columns (5) and (6) add school fixed effects, while Columns (7) and (8) include school value added computed using test scores from prior cohorts. Column (6) and (8) also control for whether the test center is located in a school or university. In both cases, results are robust to the inclusion of these center characteristics. Finally, Column (9) restricts the sample to schools with enough variation in our two main variables of interest, treatment and adversity levels.

While the baseline specification in Column (4) shows a statistically significant negative effect of noise on test scores (-0.12 SD), the magnitude of this effect decreases and becomes statistically insignificant once school fixed effects (Columns 5–6:) or school value added controls (Columns 7–8) are included (-0.015 SD). This shift likely reflects the reduction in identifying variation once comparisons are made within schools, rather than across them. Despite this attenuation, the interaction between noise exposure and adversity remains positive and significant across all specifications, suggesting that students from more adverse backgrounds are relatively less affected—or even perform slightly better—when tested under noisy conditions (0.025 SD). This result is robust to the inclusion of school value-added controls, suggesting that the observed resilience is not merely a byproduct of school-level sorting but reflects a specific interaction between the individual's developmental history and the immediate evaluative environment. Collectively, these findings support the hypothesis that early-life adversity fosters specialized skills that are selectively revealed under environmental pressure.

Figure 4 presents a visual preview of these results, showing the estimated performance gap between students assigned to noisy test centers and those tested under standard conditions.The figure is based on the preferred specification, which controls for school value added and test center characteristics (column 8, Table 2). As shown, exposure to noise is generally associated with lower scores, but for high-adversity students, it is linked to an increase of approximately 0.025 standard

deviations in test performance. Although modest in magnitude, this effect is comparable to estimates in the literature on environmental impacts on academic performance, such as those related to air pollution and heat (Park et al., 2020; Kremer et al., 2024). Subject-level breakdowns are presented in Figure 5.

The disaggregated results by subject indicate that the interaction between adversity and noise exposure is most pronounced in verbal and social domains. As shown in Table 3 and Figure 5, students from high-adversity backgrounds perform better in noisy environments in English (0.035 sd), Social Science (0.031 sd), relative to their low-adversity peers. These patterns suggest that verbal and context-based reasoning skills may be more sensitive to environmental triggers that activate adaptive responses among students accustomed to adversity. In contrast, the α in Math and Natural Science is smaller and not statistically significant, indicating that performance in these more structured domains may be less influenced by environmental variation.

Figure 6 disaggregates the effects by quartiles of adversity. The results suggest that moderate adversity can, in some contexts, lead to beneficial adaptations that influence skill development. The traditional human capital model suggests that early adversity generally has negative effects on human capital accumulation[21], as adverse conditions typically hinder the development of cognitive skills. However, introducing adaptability even in the simplest setting reveals scenarios in which early adversity can enhance outcomes. Specifically, moderate levels of adversity can promote beneficial adaptations in adaptive skills.

Taken together, these results suggest that environmental factors, such as noise, play a critical role in the expression of cognitive skills and may introduce systematic biases in standardized testing. Performance gaps shrink under adverse conditions, suggesting that conventional assessment environments do not capture the full potential of students from high-adversity backgrounds, highlighting the need for more inclusive and context-sensitive evaluation frameworks.

V. A Other measures of adversity

In addition to testing my hypothesis using the composite adversity index, I also explore three alternative proxies for adversity: (i) living in a crowded household (defined as more than six people), (ii) living in a single-parent household, and (iii) a binary indicator for whether the student was working at the time of the exam. Table 4 presents the results using each of these measures. The interaction effects for students living in crowded households and those who report working are consistent with the findings using the main adversity index, and in fact, slightly larger in magnitude. This may reflect that both working and overcrowded living conditions are more directly associated with exposure to unpredictability and competing cognitive demands, which could en-

[21] Unless parents compensate for early shocks (Cunha and Heckman, 2007).

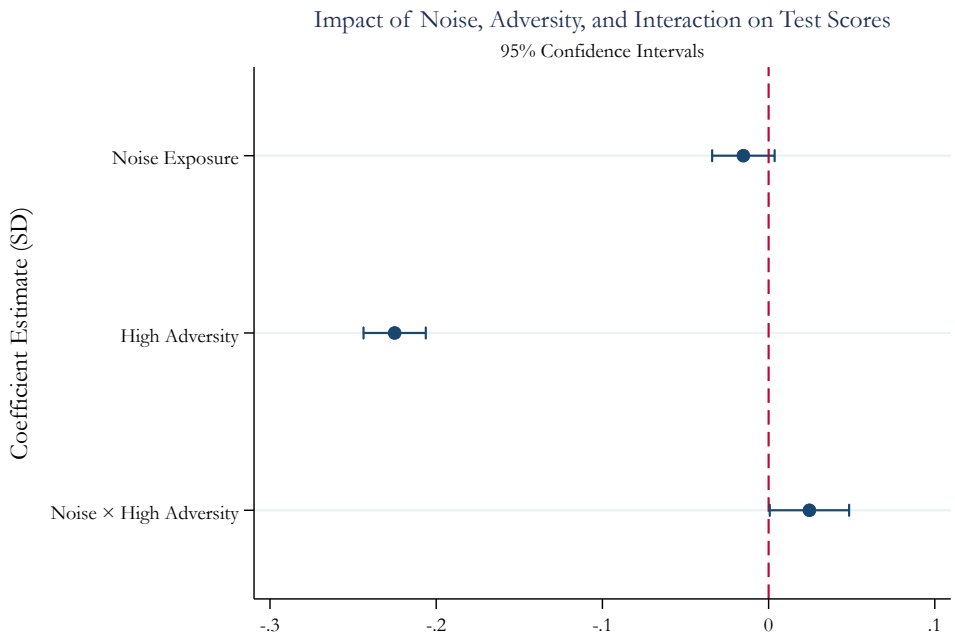
hance the development of stress-adaptive skills. In contrast, there is no evidence of a differential effect of noise for students from single-parent households, suggesting that not all dimensions of adversity trigger the same adaptive responses relevant for performance under stressful conditions.

Table 2. Noise exposure and test performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Global score (sd)								
Noisy Environment =1	-0.122***	-0.008	-0.008	-0.034***	-0.012	-0.015	-0.012	-0.015	-0.018*
	(0.006)	(0.009)	(0.009)	(0.012)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
Adversity Index = 1			-0.653***	-0.617***	-0.195***	-0.195***	-0.225***	-0.225***	-0.225***
			(0.015)	(0.021)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
α= Noisy*Adv				**0.094*** **	**0.020***	**0.019***	**0.025****	**0.025****	**0.027****
				(0.027)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)
Observations	107,734	107,734	105,776	105,758	105,758	105,758	99,794	99,794	97,313
Adjusted R-squared	0.009	0.100	0.174	0.207	0.436	0.436	0.427	0.427	0.419
Zone FE		X	X	X	X	X	X	X	X
Student Controls				X	X	X	X	X	X
School Controls					School Fixed Effects		School Value Added (t-1)		
Test Center Controls						X		X	X
Sample: schools w/variation									X
Effect				0.060	0.008	0.004	0.014	0.009	0.009
Cluster			Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center

Notes. This table reports the difference in performance between Low and High-Noise test centers. High noise pollution is defined as test-centers above 65 dB of noise. All specifications (2-9)include period-zone fixed effects. Column 9 reports the effects for a limited sample of schools with variation in adversity levels and treatment (See Table 7). Significance: ***=.01, **=.05, *=.1. Standard errors clustered at the center level.

Figure 4



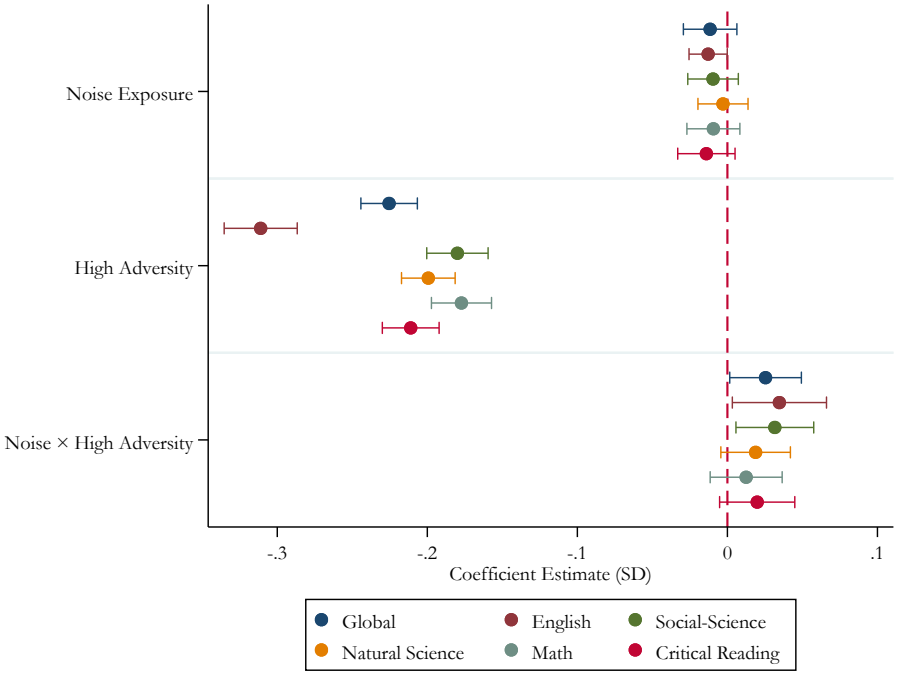
*Notes:*This figure reports the coefficients on equation 1 and column (8) of Table 2. Noisy refers to students who presented the test in centers with >65 dB.

Table 3. Noise exposure and test performance by subject

	(1) Global score	(2) English	(3) Social Science	(4) Natural Science	(5) Math	(6) Critical reading
Noisy Environment =1	-0.015	-0.013**	-0.012	-0.006	-0.014	-0.017*
	(0.010)	(0.007)	(0.009)	(0.009)	(0.009)	(0.010)
Adversity Index = 1	-0.225***	-0.311***	-0.180***	-0.199***	-0.177***	-0.211***
	(0.010)	(0.012)	(0.010)	(0.009)	(0.010)	(0.010)
α= Noisy*Adv	0.025**	0.035**	0.031**	0.018	0.011	0.019
	(0.012)	(0.016)	(0.013)	(0.012)	(0.012)	(0.013)
Observations	99,792	99,094	99,792	99,792	99,792	99,792
Adjusted R-squared	0.427	0.465	0.309	0.378	0.380	0.297

Notes. This table reports the difference in performance between Low and High-Noise test centers. High noise pollution is defined as test-centers above 65 dB of noise. All specifications include period-zone fixed effects, lagged School value-added, and test center characteristics. Significance: ***=.01, **=.05, *=.1. Standard errors clustered at the center level.

Figure 5. Effects of noise on learning outcomes by subject



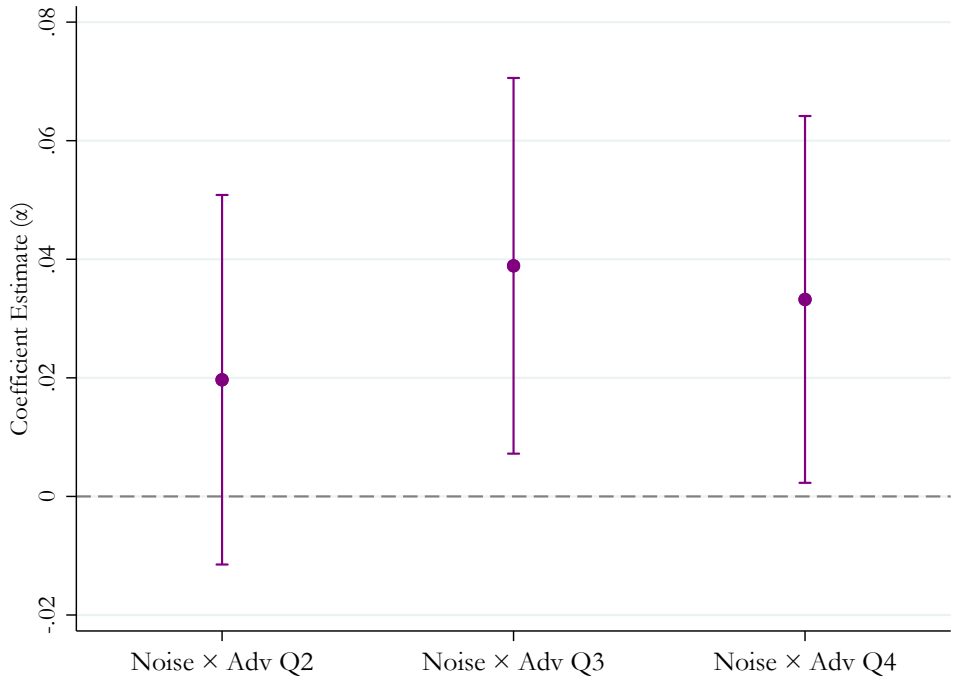
Notes: This figure reports the coefficients of Table 3, and equation 1, breakdown by subject. Noisy refers to students who presented the test in centers with >65 dB.

V. B Long-run impacts on college enrollment and human capital

To explore the impact of testing conditions on college enrollment and human capital, I estimate a measure of test center quality using an Empirical Bayes (EB) approach based on the SABER 11 performance of previous cohorts at each site. This EB measure allows me to capture the persistent, latent quality of a testing site—accounting for environmental stability and administrative efficiency—while shrinking noisy estimates from smaller centers toward the mean. I then use this measure as an instrument for college enrollment in a two-stage least squares (2SLS) framework, allowing me to identify the causal effect of higher education on graduation.

The results, presented in table 5, establish a persistent causal link between the initial testing environment and long-run human capital accumulation. The first-stage estimates in Column 1 show that Test Center Quality is a robust predictor of College Enrollment, with an F-statistic of 10.450. This confirms that the quality of the testing environment exerts a significant influence on the intensive margin of college access. In the second stage (Column 2), I estimate the impact of this enrollment on Saber Pro scores—the national exit exam taken at the conclusion of university studies. The IV estimate indicates that enrolling in a degree program leads to a substantial 0.804 standard devia-

Figure 6. Interaction Effects: Noise × Adversity Quartile

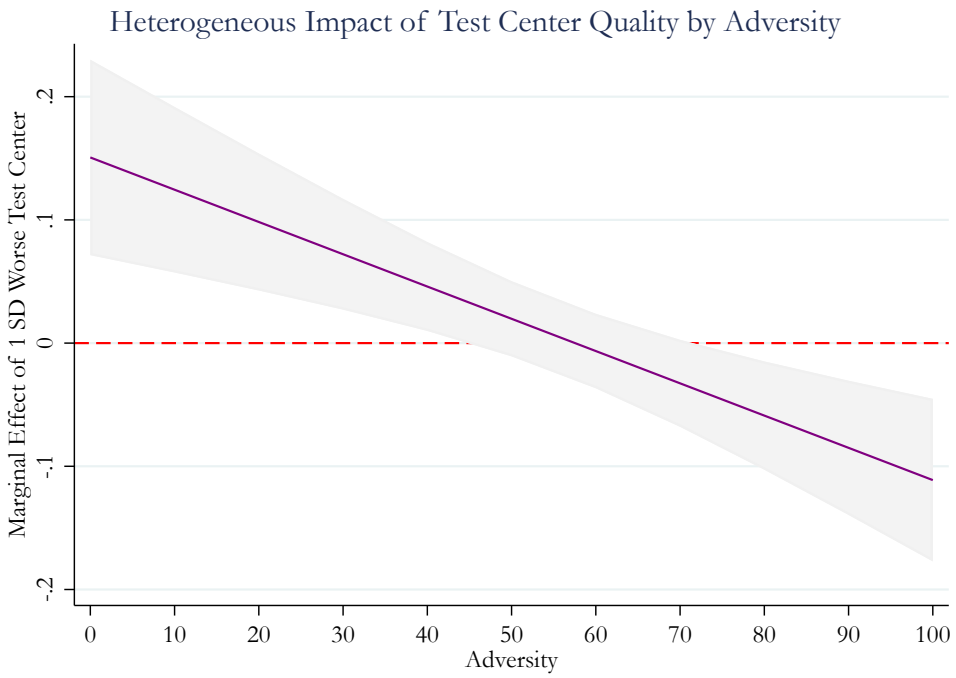


Notes: This figure reports the coefficients of the interaction effect between adversity and noisy center, breakdown by adversity quartile. The reference quartile (Q1) represents low adversity. Noisy refers to students who presented the test in centers with >65 dB.

tion increase in exit scores. These results suggest that the shocks experienced during the Saber 11 exam propagate through the educational lifecycle, ultimately influencing the probability of degree completion. Therefore, a sub-optimal testing environment does more than depress an initial score; it can fundamentally alter a student's trajectory of human capital accumulation and graduation.

Additionally, Figure 7 illustrates that the impact of being assigned to a lower-quality test center is primarily borne by students from low-adversity backgrounds. In contrast, students who experienced higher levels of adversity appear relatively unaffected by the same environmental stressors. This pattern supports the idea that students from high-adversity environments may develop adaptive skills that make them more resilient to disruptive testing conditions.

Figure 7. Marginal effect of being assigned to a 1 SD worse test center



*Notes:*This figure plots the estimated effect of being assigned to a lower-quality test center—measured by a one standard deviation decrease in empirical Bayes (EB) test center quality—across the distribution of adversity.

Table 4. Noise exposure and test performance other adversity measures

	Global Score (SD)		
Noisy Environment =1	-0.004	-0.001	-0.011
	(0.008)	(0.008)	(0.009)
More than 6 people in the HH = 1	-0.128***		
	(0.013)		
Noisy * More than 6 people in the HH	0.031*		
	(0.018)		
Single parent HH		0.024*	
		(0.013)	
Noisy * Single parent HH		-0.006	
		(0.018)	
Student works			-0.156***
			(0.011)
Noisy * Student works			0.024*
			(0.014)
Observations	97,349	97,349	99,266
Adjusted R-squared	0.423	0.422	0.421
Effect	0.027	-0.007	0.013

*Notes.*This table reports the difference in performance between Low and High-Noise test centers. High noise pollution is defined as test-centers above 65 dB of noise. All specifications include period-zone fixed effects, lagged School value-added, and test characteristics. Significance: ***=.01, **=.05, *=.1. Standard errors clustered at the center level. Adolescents who work while attending school often do so out of economic necessity, which may reflect broader household constraints. To capture this additional aspect, I construct a binary variable indicating whether a student reported working at the time of the exam. This indicator, drawn from the ICFES demographic survey, allows me to identify students who are likely navigating additional responsibilities.

Table 5. Effect of Test Center Quality on College Enrollment and Saber Pro Scores

	(1) First Stage: College Enrollment	(2) Second Stage: SaberPro (2SLS)
Test Center Quality	0.007***	
	(0.002)	
Female	-0.011***	0.083***
	(0.002)	(0.004)
Age at Exam	-0.021***	-0.014**
	(0.001)	(0.006)
LCI	0.012***	0.004
	(0.000)	(0.003)
Enroll in a HE degree		0.804***
		(0.302)
F-statistic	10.450	
Prob > F	0.0012	
Observations	150,780	150,780

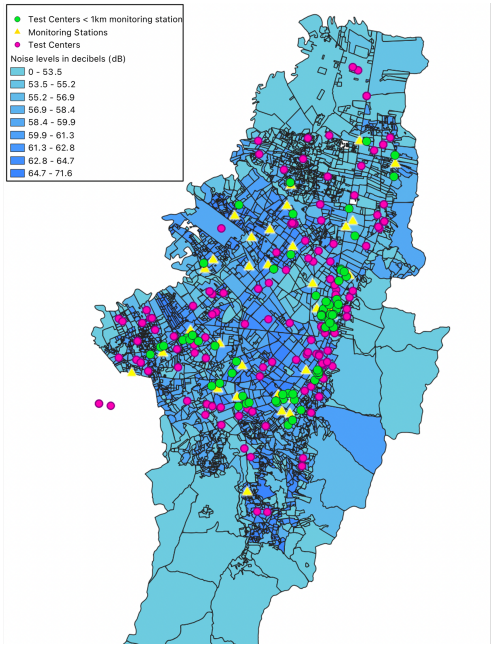
Notes. This table presents instrumental variable (IV) estimates of the impact of college enrollment on Saber Pro exit exam scores. Column 1 reports the first-stage relationship between the quasi-random assignment to test center quality and the probability of enrolling in a higher education (HE) degree. Column 2 reports the second-stage (2SLS) estimates, where enrollment is instrumented by test center assignment. All specifications include controls for gender, age at the time of the Saber 11 exam, and the Living Conditions Index (LCI). Standard errors are clustered at the test-center level to account for within-site correlation. * p ¡ 0.1, ** p ¡ 0.05, *** p ¡ 0.01.

VI Robustness Checks

i. Other specifications

To assess the robustness of my findings, I conducted several additional exercises. First, Table 6 re-estimates the main results using only test centers located within 1 kilometer of a noise monitoring station, to ensure more accurate measurement of environmental exposure (See Figure 8). The interaction term (α) in Equation 1 remains positive and statistically significant (0.03 standard deviations), although the significance level is slightly lower than in the full sample. Notably, the coefficient on the main noise variable flips sign in this subsample, suggesting that attenuation bias due to measurement error in noise exposure may be influencing the estimates in the broader sample.

Figure 8. Monitoring Stations and Test centers ($<$ 1Km)



Notes: Source: constructed based on the Environmental Noise Monitoring Network (RMRAB).

Table 6. Effects of noise on learning outcomes (Only test centers $<$ 1km)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
						Global score					
Noisy Environment =1	-0.211***	0.001	0.005	-0.025*	-0.025*	0.001	0.002	-0.003	-0.002	0.005	0.006
	(0.009)	(0.011)	(0.010)	(0.015)	(0.015)	(0.011)	(0.011)	(0.013)	(0.013)	(0.011)	(0.011)
Adversity Index = 1			-0.661***	-0.727***	-0.629***	-0.213***	-0.213***	-0.247***	-0.246***	-0.242***	-0.242***
			(0.021)	(0.031)	(0.028)	(0.014)	(0.014)	(0.017)	(0.017)	(0.014)	(0.014)
= Noisy*Adv				**0.123*** **	**0.113*** **	0.027	0.026	**0.044***	**0.044***	**0.035***	**0.035***
				(0.040)	(0.037)	(0.018)	(0.018)	(0.023)	(0.023)	(0.019)	(0.019)
Observations	48,186	48,186	47,286	47,286	47,280	47,280	47,280	44,145	44,145	44,463	44,463
Adjusted R-squared	0.012	0.103	0.180	0.181	0.214	0.435	0.435	0.349	0.349	0.427	0.427
Zone FE		X	X	X	X	X	X	X	X	X	X
Student Controls					X	X	X	X	X	X	X
School Controls						FEs	FEs	X + Score p-tile		Value Added	
Test Center Controls							X		X		X
Effect				0.066	0.060	0.008	0.004	0.018	0.013	0.014	0.009
Cluster			Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center

Notes. This table reports the difference in performance between Low and High-Noise test centers, keeping the 98 centers below 1 KM of distance from a monitoring station (Figure 12). High noise pollution is defined as test-centers above 65 dB of noise. All specifications (2-11)include period-zone fixed effects. Significance: ***=.01, **=.05, *=.1. Standard errors clustered at the center level.

Second, I restrict the sample to schools that exhibit sufficient within-school variation in both adversity and noise exposure. Table 7 presents the results for 1,789 out of 1,977 schools where the proportion of high-adversity students falls between 10% and 80%. In this restricted sample, the coefficient on the noise variable becomes statistically significant, while the interaction term remains positive and significant. These results contrast with the main sample, where the noise coefficient was not statistically significant. This suggests that the lack of significance in the full sample is likely due to limited within-school variation rather than bias in the estimation. The findings are consistent with the simulation exercise, which showed that attenuation of the noise effect under school fixed effects is expected under random assignment with limited identifying variation. Together, these results support the interpretation that the empirical strategy is valid and that the observed patterns are not driven by selection or omitted variables.

Finally, Table 8 shows that the results are robust to different choices of clustering level. The table replicates the preferred specification from Table 2, adjusting the clustering from test center to zone-by-period. Across all specifications, the estimated interaction effect remains positive and comparable in magnitude, suggesting that the results are not sensitive to the inference strategy.

ii. Simulation

To evaluate the robustness of my identification strategy, I assess how the estimated treatment effect behaves when school fixed effects are included in the model. While my baseline specification finds a significant negative effect of noise exposure on student performance, the coefficient becomes smaller and statistically insignificant when school fixed effects are added. This attenuation could suggest either a loss of identifying variation or a concern with the specification. To test this, I run a simulation that randomizes noise exposure within schools—preserving the structure of the identifying variation used in the school fixed effects specification—while holding all covariates constant.

Table 7. Effects of noise on learning outcomes (Only using schools with variation of adversity and noise)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Global score										
Noisy Environment =1	-0.134***	-0.011	-0.010	-0.039***	-0.040***	-0.014	-0.017*	-0.021**	-0.026**	-0.014	-0.018*
	(0.006)	(0.009)	(0.009)	(0.012)	(0.012)	(0.009)	(0.009)	(0.010)	(0.011)	(0.009)	(0.010)
Adversity Index = 1			-0.645***	-0.710***	-0.617***	-0.195***	-0.195***	-0.229***	-0.229***	-0.225***	-0.225***
			(0.014)	(0.023)	(0.021)	(0.010)	(0.010)	(0.012)	(0.012)	(0.010)	(0.010)
= Noisy*Adv				**0.112*** **	**0.107*** **	**0.021***	**0.021***	**0.040*** **	**0.040*** **	**0.028****	**0.027****
				(0.028)	(0.026)	(0.012)	(0.012)	(0.015)	(0.015)	(0.012)	(0.012)
Observations	105,096	105,096	103,186	103,186	103,169	103,169	103,169	96,808	96,808	97,313	97,313
Adjusted R-squared	0.006	0.092	0.166	0.167	0.199	0.427	0.427	0.342	0.342	0.419	0.419
Zone FE		X	X	X	X	X	X	X	X	X	X
Student Controls					X	X	X	X	X	X	X
School Controls						FEs	FEs	X + Score p-tile		Value Added	
Test Center Controls							X		X		X
Effect				0.073	0.067	0.008	0.004	0.019	0.014	0.014	0.009
Cluster			Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center	Test-Center

Notes. This table reports the difference in performance between Low and High-Noise test centers for the sample of 1,789 out of 1,977 schools with enough adversity variation. Adversity variation is defined as having a proportion of students from an adverse background between 10% and 80%. High noise pollution is defined as test-centers above 65 dB of noise. All specifications (2-8)include period-zone fixed effects. Significance: ***=.01, **=.05, *=.1. Standard errors clustered at the center level.

Table 8. Effects of noise on learning outcomes (Different Clusters)

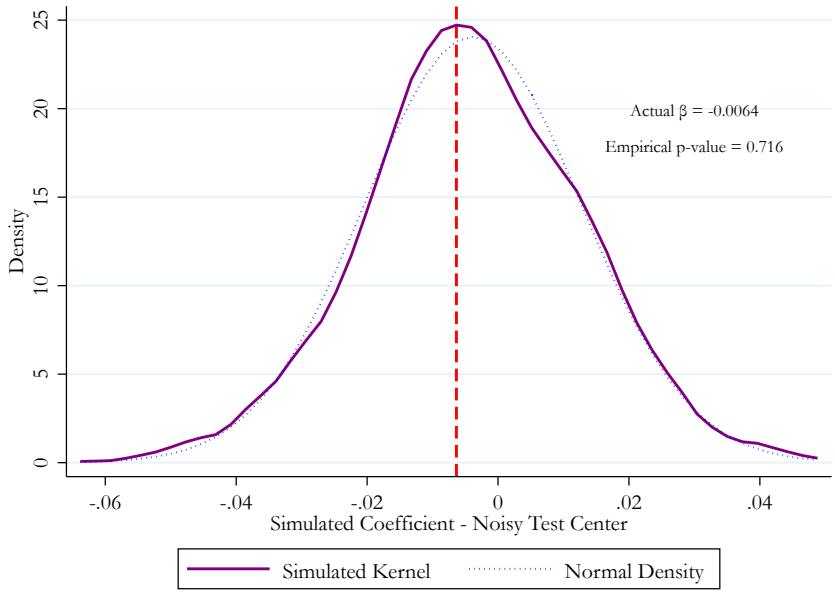
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Global Score											
Noisy Environment =1	-0.122***	-0.008	-0.008	-0.008	-0.008	-0.008	-0.033***	-0.033**	-0.010	-0.010	-0.012	-0.012
	(0.006)	(0.007)	(0.009)	(0.005)	(0.009)	(0.005)	(0.012)	(0.013)	(0.009)	(0.009)	(0.009)	(0.008)
Adversity Index = 1					-0.653***	-0.653***	-0.710***	-0.710***	-0.231***	-0.231***	-0.195***	-0.195***
					(0.015)	(0.047)	(0.023)	(0.051)	(0.010)	(0.009)	(0.010)	(0.010)
α= Noisy*Adv							0.099***	0.099***	0.019*	0.019*	0.020*	0.020*
							(0.029)	(0.036)	(0.012)	(0.011)	(0.011)	(0.010)
Observations	107,734	107,734	107,734	107,734	105,776	105,776	105,776	105,776	105,665	105,665	105,647	105,647
Adjusted R-squared	0.009	0.100	0.100	0.100	0.174	0.174	0.174	0.174	0.421	0.421	0.436	0.436
Mean control												
Year	All	All	All	All	All	All	All	All	All	All	All	All
Controls											X	X
School controls												
School FE									X	X	X	X
Zone FE		X	X	X	X	X	X	X	X	X	X	X
Cluster			Center	Zone-period	Center	Zone-period	Center	Zone-period	Center	Zone-period	Center	Zone-period

Notes. This table reports the difference in performance between Low and High-Noise test centers. High noise pollution is defined as test-centers above 65 dB of noise. All specifications (2-8)include period-zone fixed effects. Significance: ***=.01, **=.05, *=.1. Standard errors clustered at the center level.

For each iteration, I reassign treatment randomly within schools to match the actual number of treated students in each school, estimate the treatment effect using the same model (including school fixed effects), and record the coefficient. The distribution of simulated coefficients is centered around zero, and the observed estimate from the real data lies well within the simulated range.

Figure 9 shows the distribution of simulated coefficients alongside the actual estimate (red line). The empirical p-value of 0.716 indicates that my observed coefficient is typical under random assignment, suggesting that the loss of statistical significance in the main specification is consistent with the limited within-school variation in treatment, rather than evidence of bias.

Figure 9. Simulated vs. Actual Treatment Effect
Kernel and Normal Density



Notes: This figure shows the simulated distribution of treatment effects coefficients under random assignment of noise exposure within schools, along with the actual estimate (red line). The p-value of 0.716 indicates that my observed coefficient is typical under random assignment, suggesting that the loss of statistical significance in the main specification is driven by limited within-school variation in treatment, rather than evidence of bias.

VII Simple Theoretical Framework

To guide my hypotheses, I propose a theoretical framework that includes stress-adaptive skills (A_t) in a traditional human capital model to formalize that early adversity (E_t) can, in some contexts, lead to beneficial adaptations that influence skill development. The framework highlights the context-dependent nature of adaptive skills, which may only emerge under specific conditions or testing environments (X)İt also sheds light on how this dependence can lead to measurement bias

in the traditional methods used to evaluate skills.

I consider a standard model of human capital formation, where childhood is divided into multiple developmental periods t $\in$ [0,1] [22], and adult outcomes (Y) are determined by traditional and adaptive skills children gain at the end of childhood. Traditional skills include academic, problem-solving abilities as well as non-cognitive skills, while adaptive skills are stress-response capabilities shaped by early adversity (E_t) and influenced by adaptation[23]. Both skill types are shaped by parental investments (I_t), early adversity (E_t), and idiosyncratic shocks (ε_t).

The development of traditional (C) skills is defined as $C_{t+1} = f_C(C_t, I_{t+1}, E_{t+1}, \varepsilon_{C,t+1})$, while the development of adaptive skills follows $A_{t+1} = f_A(A_t, I_{t+1}, E_{t+1}, \varepsilon_{A,t+1})$. Importantly, adaptive skills emerge under specific testing conditions (X) and are linked to an individual's ability to respond to stress or adversity.

Adult outcomes, such as test scores (Y), depend on the levels of both skill types at the end of childhood (C_1 and A_1) and contextual factors (X) that affect the returns to these skills. This relationship is captured by $Y = g(C_1, A_1, X)$, where g is continuously differentiable and increasing in both skill dimensions.

The traditional human capital model suggests that early adversity generally has negative effects on human capital accumulation[24], as adverse conditions typically hinder the development of cognitive skills, represented by $\frac{\partial Y}{\partial E_t} < 0$ through $\frac{\partial C_t}{\partial E_t} < 0$. However, introducing adaptability even in the most simple setting reveals scenarios in which early adversity can enhance outcomes. Specifically, adversity can promote beneficial adaptations in adaptive skills, reflected by $\frac{\partial A_t}{\partial E_t} > 0$. As a result, the overall effect of early adversity on outcomes is captured by:

$$\frac{\partial Y}{\partial E_t} = \frac{\partial C_t}{\partial E_t} + \frac{\partial A_t}{\partial E_t}.$$

This framework provides a more nuanced perspective on how adversity influences skill development and generates several testable empirical predictions. Moderate levels of adversity are expected to enhance adaptive skills, whereas extreme adversity limits overall skill formation. In high-stress environments, the returns to adaptive skills (A_1) are amplified, making them a critical determinant of outcomes ($\frac{\partial Y}{\partial A_1} > 0$ under high-stress conditions). In contrast, the effect on traditional skills tend to dominate in low-stress settings, where the returns to C_1 exceed those of A_1 ($\frac{\partial Y}{\partial C_1} > \frac{\partial Y}{\partial A_1}$ under low-stress conditions).

[22] In this case, I limit to two periods for simplicity. However, the model could be extended to multiple periods following (Cunha and Heckman, 2007)

[23] For simplicity, I do not differentiate between cognitive and non-cognitive skills, because stress-adaptive skills can be both cognitive or non-cognitive skills.

[24] Unless parents compensate for early shocks (Cunha and Heckman, 2007).

VIII Conclusion

The mismeasurement of abilities limits our understanding of human capital formation and can reinforce socioeconomic inequalities by misallocating educational and occupational opportunities (Ebenstein et al., 2016), which in turn may slow structural transformation. While this study is set in Bogotá, Colombia, the issues it raises are relevant for education systems more broadly, many of which rely on standardized assessments administered in controlled, uniform environments that may not capture the abilities of students who have developed skills in more unpredictable settings.

This paper provides new evidence on how environmental conditions influence the expression of cognitive skills, particularly among students from disadvantaged backgrounds. Using quasi-random variation in exposure to noise during Colombia's national high school exit exam, I show that while noise generally reduces test performance, students with high levels of childhood adversity exhibit adaptability and even perform better. The interaction between adversity and noise exposure is positive and statistically significant, suggesting that conventional testing environments may systematically underestimate the abilities of disadvantaged students.

The main implication of this paper is that governments/organizations around the world aiming to measure individual cognitive ability at best should re-weight the exam marks based on test center conditions. Although high-stakes exams play an important role in enabling comparisons across students from schools with varying grading standards, the findings indicate that these assessments may offer a somewhat noisy signal of student quality. The results offer valuable insight into the importance of accounting for testing conditions when interpreting standardized scores, especially for students from vulnerable backgrounds. Despite this, one limitation in this analysis is that it cannot disentangle whether the observed effects are driven by distraction-related impairment or the activation of stress-adaptive skills among students from high-adversity backgrounds, highlighting the need for the controlled experimental component of this study.

References

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham (2020) "The Light and the Heat: Productivity Co-benefits of Energy-saving Technology," *The Review of Economics and Statistics*, 102 (4), 779–792, 10.1162/rest_a_00832.

Almond, Douglas and Janet Currie (2011) "Human Capital Development Before Age Five," in Ashenfelter, Orley and David Card eds. *Handbook of Labor Economics*, 4b: Elsevier.

Ashraf, Nava, Abhijit Banerjee, and Vesall Nourania (2020) "Learning to Teach by Learning to Learn," November, Job Market Paper.

Bailey, Charles E (2007) "Cognitive Accuracy and Intelligent Executive Function in the Brain and in Business," *Annals of the New York Academy of Sciences*, 1118, 122–141.

Baldiga, Katherine (2014) "Gender Differences in Willingness to Guess," *Management Science*, 60 (2), 434–448, 10.1287/mnsc.2013.1776.

Banerjee, Abhijit V., Swati Bhattacharjee, Raghabendra Chattopadhyay, Esther Duflo, Alejandro J. Ganimian, Kailash Rajah, and Elizabeth S. Spelke (2025) "Children's arithmetic skills do not transfer between applied and academic mathematics," *Nature*, 10.1038/s41586-024-08502-w.

Bassi, Vittorio and Aisha Nansamba (2021) "Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda," *The Economic Journal*, 132 (642), 471–511, 10.1093/ej/ueab071.

Bau, Natalie, Martin Rotemberg, Manisha Shah, and Bryce Steinberg (2020) "Human Capital Investment in the Presence of Child Labor," NBER Working Paper 27241, National Bureau of Economic Research, 10.3386/w27241, Last revised: 11 Jun 2025.

Bessone, Pedro, Gautam Rao, Frank Schilbach, Heather Schofield, and Mattie Toma (2021) "The Economic Consequences of Increasing Sleep Among the Urban Poor," *The Quarterly Journal of Economics*, 136 (3), 1887–1941, 10.1093/qje/qjab006.

Blair, Clancy and C. Cybele Raver (2014) "Closing the achievement gap through modification of neurocognitive and neuroendocrine function: Results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten," *PLoS ONE*, 9, e112393, 10.1371/journal.pone.0112393.

Borella, Erika, Barbara Carretti, and Santiago Pelegrina (2010) "The Specific Role of Inhibition in Reading Comprehension in Good and Poor Comprehenders," *Journal of Learning Disabilities*, 43 (6), 541–552.

Bradler, Christian, Susanne Neckermann, and Arne J. Warnke (2019) "Incentivizing Creativity: A Large-Scale Experiment with Performance Bonuses and Gifts," *Journal of Labor Economics*, 37 (3), 793–851.

Brienza, Justin P. and Igor Grossmann (2017) "Social class and wise reasoning about interpersonal conflicts across regions, persons and situations," *Proceedings of the Royal Society B*, 284 (1868), 20171870, 10.1098/rspb.2017.1870.

Center on the Developing Child (2012) "Executive Function (InBrief)," https://developingchild.harvard.edu/resources/inbrief-executive-function/.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell (2016) "Particulate Pollution and the Productivity of Pear Packers," *American Economic Journal: Economic Policy*, 8 (3), 141–169, 10.1257/pol.20150085.

——— (2019) "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China," *American Economic Journal: Applied Economics*, 11 (1), 151–172, 10.1257/app.20160436.

Cunha, F. and J. J. Heckman (2007) "The technology of skill formation," *American Economic Review*, 97 (2), 31–47.

Dahlman, Sandra, Pernilla Bäckström, Gunilla Bohlin, and Örjan Frans (2013) "Cognitive abilities of street children: Low-SES Bolivian boys with and without experience of living in the street," *Child Neuropsychology*, 19, 540–556, 10.1080/09297049.2012.731499.

Dean, Emma Boswell, Frank Schilbach, and Heather Schofield (2017) "Poverty and Cognitive Function," https://example.com, Unpublished Manuscript.

Dean, Joshua T. (2024) "Noise, Cognitive Function, and Worker Productivity," *American Economic Journal: Applied Economics*, 16 (4), 322–360, 10.1257/app.20220532.

Deming, David J. (2017) "The Growing Importance of Social Skills in the Labor Market," *The Quarterly Journal of Economics*, 132 (4), 1593–1640, 10.1093/qje/qjx022.

Dittmann, Andreas G., Nicole M. Stephens, and Sarah S. Townsend (2020) "Achievement is not class-neutral: Working together benefits people from working-class contexts," *Journal of Personality and Social Psychology*, 10.1037/pspa0000194.

Doyle, Orla (2020) "The first 2,000 days and child skills," *Journal of Political Economy*, 128 (6).

Duncan, Greg J, Chantelle J Dowsett, Amy Claessens et al. (2007) "School Readiness and Later Achievement," *Developmental Psychology*, 43 (6), 1428–1446.

Duncan, Greg, Ariel Kalil, Magne Mogstad, and Mari Rege (2023) "Investing in early childhood development in preschool and at home," in *Handbook of the Economics of Education*, 6, 10.1016/bs.hesedu.2022.11.005.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth (2016) "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution," *American Economic Journal: Applied Economics*, 8 (4), 36–65, 10.1257/app.20150213.

Ellis, Bruce J. and Marco Del Giudice (2019) "Developmental adaptation to stress: An evolutionary perspective," *Annual Review of Psychology*, 70, 111–139, 10.1146/annurev-psych-122216-011732.

Evans, Gary W., David Li, and Sarah S. Whipple (2013) "Cumulative risk and child development," *Psychological Bulletin*, 139, 1342–1396, 10.1037/a0031808.

Falk, Richard F and Norman B Miller (1998) "The reflexive self: A sociological perspective," *Social Psychology Quarterly*, 61 (1), 68–83, 10.2307/2787100.

Faria, Carolina de Azevedo, Heloisa Veiga Dias Alves, and Helenice Charchat-Fichman (2015) "The most frequently used tests for assessing executive functions in aging," *Dementia & Neuropsychologia*, 9 (2), 149–155, 10.1590/1980-57642015DN92000009.

Fehr, Daniel, Günther Fink, and B. Kelsey Jack (2022) "Poor and Rational: Decision-Making under Scarcity," *Journal of Political Economy*, 130 (11), 10.1086/720466.

Fink, Günther, Atheendar S. Venkataramani, and Arianna Zanolini (2021) "Early life adversity, biological adaptation, and human capital: evidence from an interrupted malaria control program in Zambia," *Journal of Health Economics*, 80, 10.1016/j.jhealeco.2021.102532.

Frankenhuis, Willem E. and Carolina de Weerth (2013) "Does early-life exposure to stress shape or impair cognition?" *Current Directions in Psychological Science*, 22, 407–412, 10.1177/0963721413484324.

Fry, Catherine E. (2018) *Executive Functions, Creativity, and Mental Health in Homeless Young People: Implications for Housing Outcome* Ph.D. dissertation, Cardiff University.

Graff Zivin, Joshua and Matthew Neidell (2012) "The Impact of Pollution on Worker Productivity," *American Economic Review*, 102 (7), 3652–3673, 10.1257/aer.102.7.3652.

Graff Zivin, Joshua, Yingquan Song, Qu Tang, and Peng Zhang (2020) "Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China," *Journal of Environmental Economics and Management*, 104, 102365, https://doi.org/10.1016/j.jeem.2020.102365.

Guilford, Joy Paul (1967) *The Nature of Human Intelligence*: McGraw-Hill.

Heckman, James J. and Tim Kautz (2013) "Fostering and measuring skills: Interventions that improve character and cognition," Working Paper 19656, National Bureau of Economic Research, https://www.nber.org/papers/w19656.

Jackson, C. Kirabo (2018) "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes," *Journal of Political Economy*, 126 (5), 2072–2107, 10.1086/699018.

Jihao Hong Justin, Li Xuan, Lei Victor (2025) "Testing Away from One's Own School: Exam Location and Performance in High-Stakes Exams," June, http://www.edworkingpapers.com/ai25-1220.

Kim, Ji-Hye, Byung-Kook Lee, and Tae-Heon Kim (2020) "Effects of self-reflection–focused career course on career adaptability," *Journal of Career Development*, 47 (6), 671–685, 10.1177/0894845319840432.

Kraus, Michael W., Paul K. Piff, Rodolfo Mendoza-Denton, Michelle L. Rheinschmidt, and Dacher Keltner (2012) "Social class, solipsism, and contextualism: How the rich are different from the poor," *Psychological Review*, 119, 546–572, 10.1037/a0028756.

Kremer, Michael, Mauricio Romero, and Santiago Saavedra (2024) "Do Air Filters Improve Student Learning? Experimental Evidence from Colombia," Working Paper.

McLaughlin, Katie A., Margaret A. Sheridan, and Heather K. Lambert (2014) "Childhood adversity and neural development: Deprivation and threat as distinct dimensions of early experience," *Neuroscience & Biobehavioral Reviews*, 47, 578–591, 10.1016/j.neubiorev.2014.10.012.

Mittal, Chiraag, Vladas Griskevicius, Jeffry A. Simpson, Sooyeon Sung, and Erika S. Young (2015) "Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function," *Journal of Personality and Social Psychology*, 109, 604–621, 10.1037/pspi0000028.

Mogstad, Magne and Gaute Torsvik (2023) "Family background, neighborhoods, and intergenerational mobility," in *Handbook of the Economics of the Family*, 1, 327–387.

Neilson, Christopher A. (2021) "Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students," Working Paper 2021-48, Princeton University, Department of Economics, https://christopherneilson.github.io/work/documents/Neilson_JMP/Neilson_SEPVouchers2021.pdf, R&R at Econometrica.

Park, Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith (2020) "Heat and Learning," *American Economic Journal: Economic Policy*, 12 (2), 306–339, 10.1257/pol.20180612.

Raven, John C. (1936) "Mental Tests Used in Genetic Studies: The Performance of Related Individuals on Tests Mainly Educative and Mainly Reproductive," Master's thesis, University of London.

Rosati, Alexandra G, Melissa Emery Thompson, Rebeca Atencia, and Joshua W Buckholtz (2023) "Distinct developmental trajectories for risky and impulsive decision-making in chimpanzees," *Journal of Experimental Psychology: General*.

Schilbach, Frank (2019) "Alcohol and Self-Control: A Field Experiment in India," *American Economic Review*, 109 (4), 1290–1322, 10.1257/aer.20170458.

Sternberg, Robert J (2021) "Adaptive Intelligence: Its Nature and Implications for Education," *Educ. Sci*, 11, 823, 10.3390/educsci11120823.

Torrance, Ellis Paul (1968) *Torrance Tests of Creative Thinking*, Princeton, NJ: Personnel.

——— (1998) *Torrance Tests of Creative Thinking: Norms-technical manual: Figural (streamlined) forms A & B*, Bensenville, IL: Scholastic Testing Service.

Wantchekon, Leonard and Sally Zhang (2024) "Empowered by Adversity: Economic Shocks and Noncognitive Skill Development in Ethiopian Youth," Working Paper 33305, National Bureau of Economic Research, Cambridge, MA, http://www.nber.org/papers/w33305.

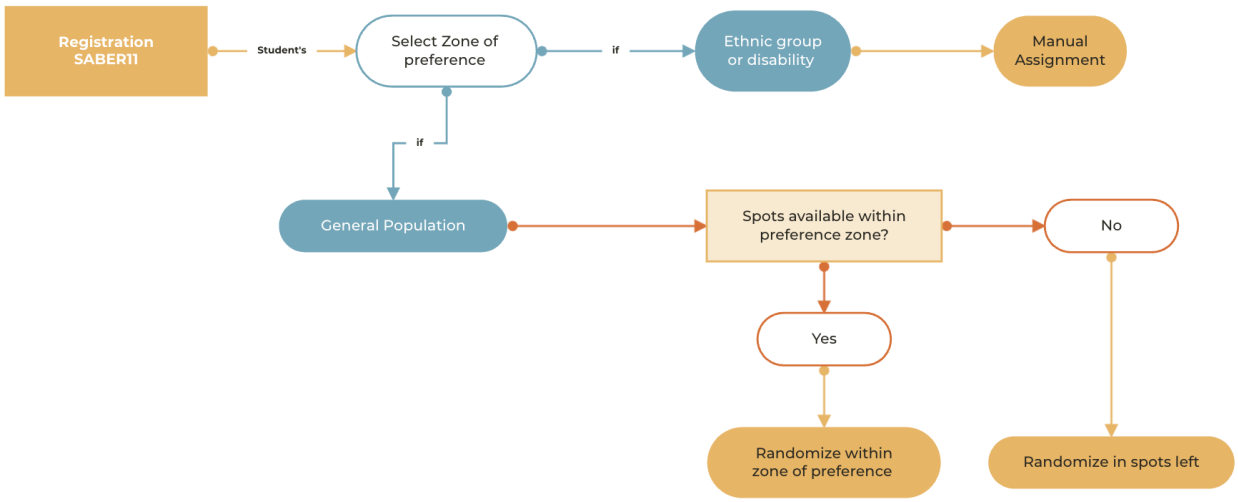
World Health Organization (1999) *Guidelines for Community Noise*, Geneva: World Health Organization, https://www.who.int/publications/i/item/a68672, Accessed 10 Dec 2023.

Young, Elizabeth S., Vladas Griskevicius, Jeffry A. Simpson, Theodore E. Waters, and Chiraag Mittal (2018) "Can an unpredictable childhood environment enhance working memory? Testing the sensitized-specialization hypothesis," *Journal of Personality and Social Psychology*, 114, 891–908, 10.1037/pspi0000124.

IX Appendix

IX. A Test Center Assignment

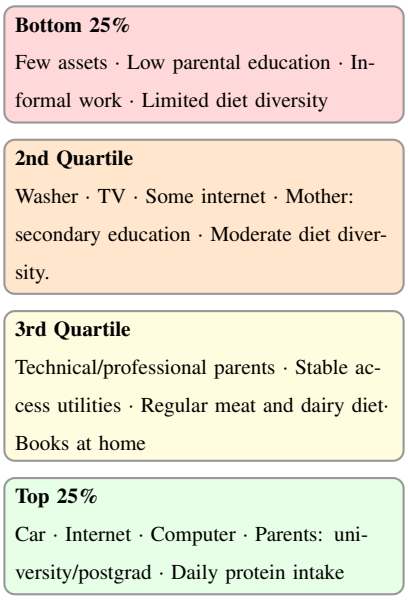
Figure 10. Test Center Assignment



IX. B Living Conditions Index (LCI)

To capture the socioeconomic baseline of the participants, we utilize the Living Conditions Index (LCI), a multidimensional proxy for household welfare developed by LCI. This index moves beyond simple income reporting by aggregating three distinct pillars of a student's home environment: Household Assets, which tracks the presence of technology (computers, internet), transportation, and essential appliances; Parental Human Capital, accounting for both educational attainment and occupational status; and Nutritional Consumption, which serves as a proxy for financial stability by measuring the frequency of protein and dairy intake.

Figure 11. Living Conditions Index quartiles



IX. C Acustic Model (MER)

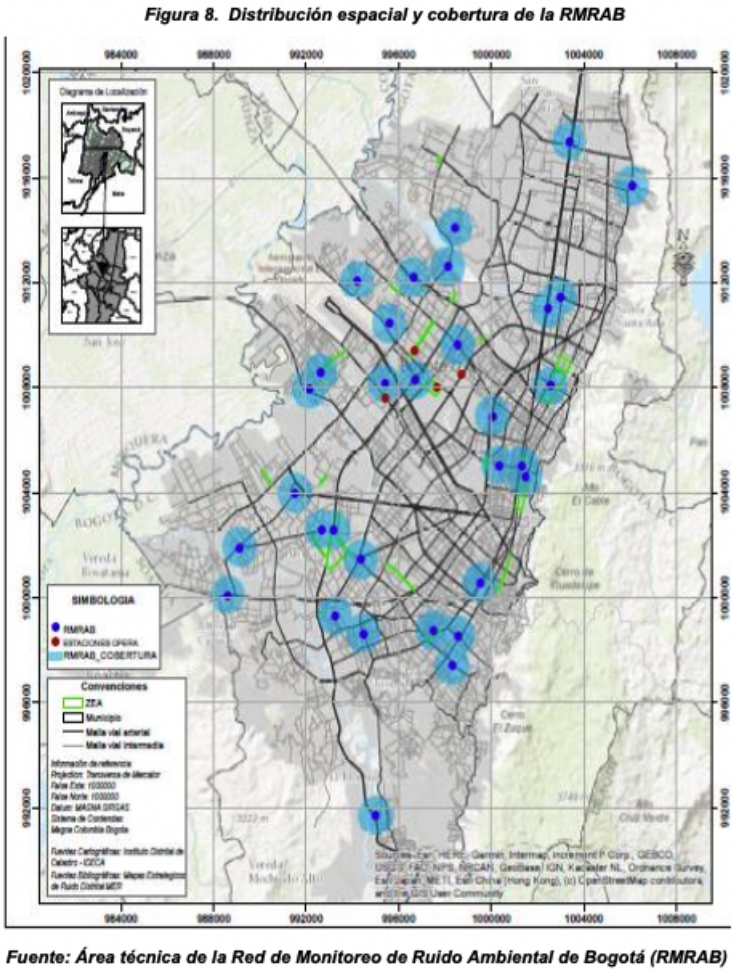
Using data from a real-time acoustic monitoring system composed of 36 stations distributed across the city (Figure 12). The MER are built using a physics-based acoustic propagation model that estimates noise exposure by simulating how sound travels and interacts with the urban environment. Key components include:

- **Sound Sources:**
  - Road traffic (disaggregated by vehicle type: cars, motorcycles, buses, trucks)
  - Air traffic (airport corridors)
  - Commercial and industrial activity
  - TransMilenio (mass transit system)
- **Environmental Conditions:**
  - Temperature, humidity, wind speed and direction (affect sound propagation)
- **Urban Geography:**
  - Building height and density
  - Land use types (residential, commercial, industrial)

  - Road network structure and elevation

- **Propagation Modeling:**
  - **ISO 9613-2 standard:** models how sound attenuates over distance
  - **CNOSSOS-EU:** models road traffic noise by vehicle type

Figure 12. Monitoring Stations



Notes: The blue dots and clusters correspond to monitoring stations. Source: RMRAB.

IX. D Noise and adversity variation across zones

Figure 13. Noise by Zone: Mean ± 1 SD

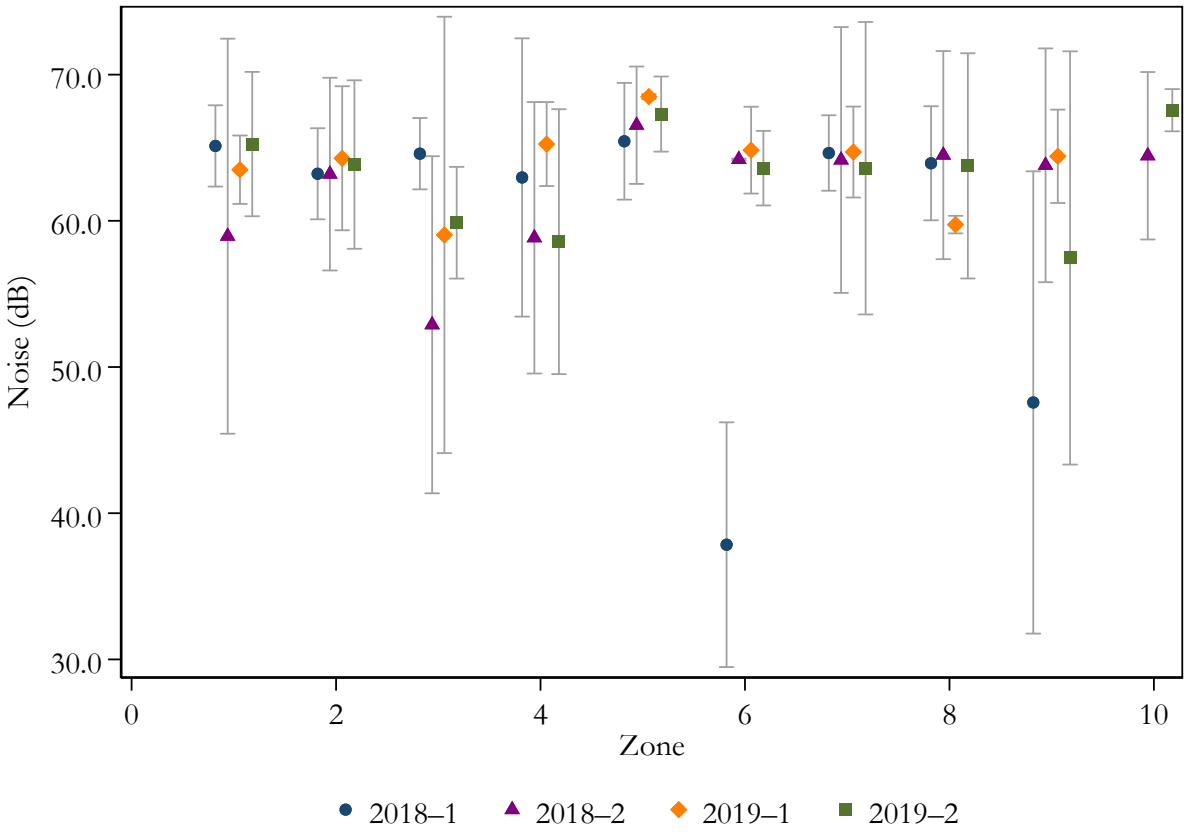
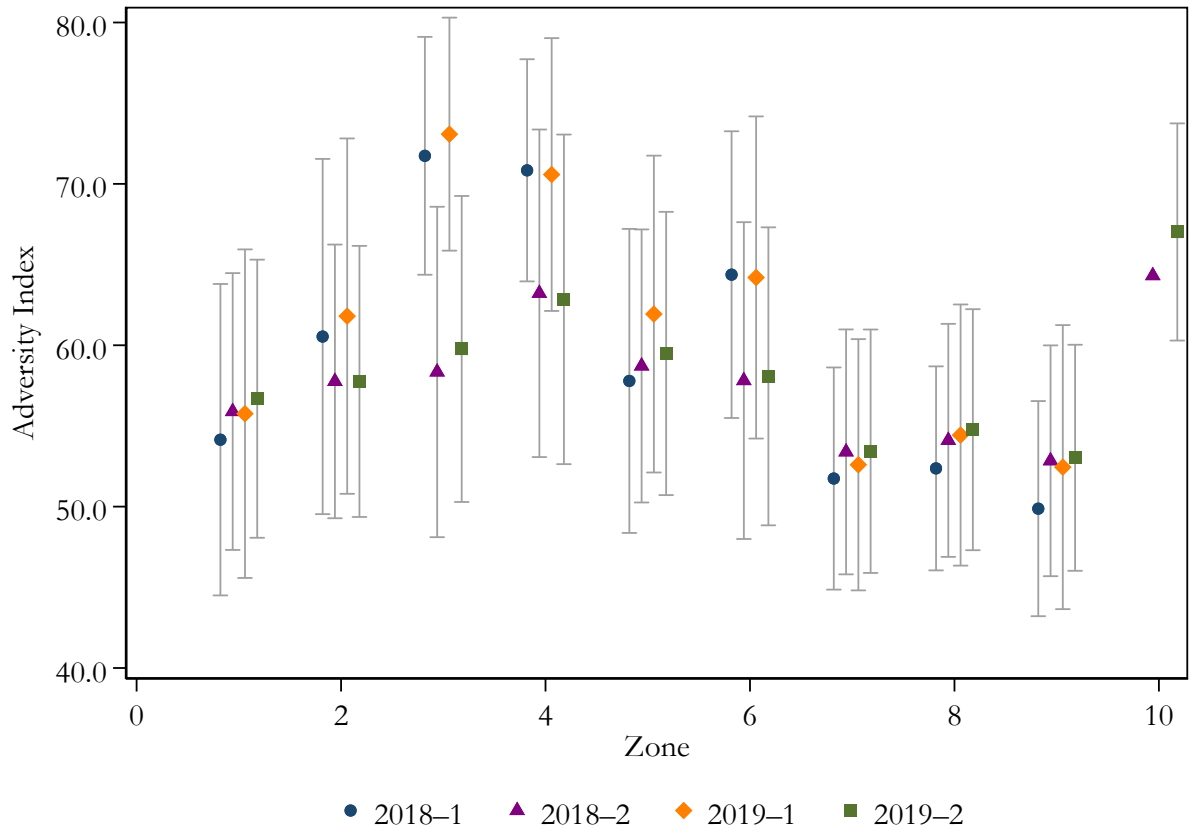


Figure 14. Adversity by Zone: Mean ± 1 SD

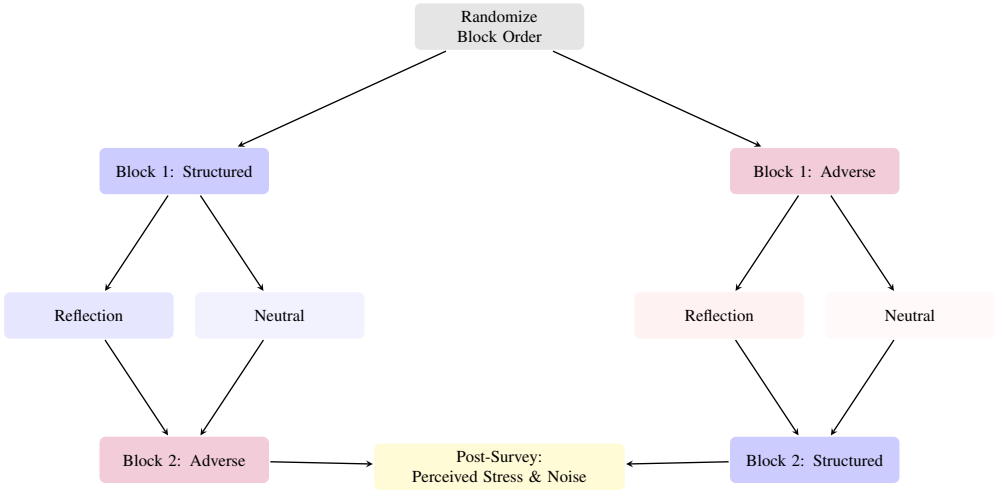


IX. E Lab experiment experimental design

To identify the mechanisms behind the quasi-experimental findings, in the second part of this project, I propose to run a lab experiment in which participants complete a battery of cognitive and stress-adaptive skills tests, including attention shifting, working memory updating, creativity, resilience, and adaptive intelligence. Each participant will be assessed under both a structured and an adverse environment. In the structured condition, tasks are completed in a quiet, predictable setting with stable rules and orderly sequencing. In the adverse condition, participants face calibrated disruptions— background noise, screen lags, and unpredictable time reminders—designed to mimic real-world instability. The order of exposure to these environments is individually randomized, allowing for clean identification of whether and how performance varies across contexts within the same individual. Randomization will be stratified by socioeconomic

status and childhood adversity (ACE scores)[25].

Figure 15. Experimental Design



In addition, all participants will be cross-randomized into a self-reflection intervention, in which they pause midway through the test to briefly evaluate the strategy they used, the challenges they faced, and one adjustment they could try (Figure 16).[26] This design allows us to test both how adversity shapes the expression of stress-adaptive skills and whether adaptability can be fostered through a simple, low-cost intervention embedded directly into the testing process.

Figure 16. Self-Reflection vs. Neutral Prompt Conditions

Self-Reflection

Prompt: *"Think about a small challenge you faced while completing the test. Describe the strategy you used, what made it difficult, and one adjustment you could try next time."*

Time: 2–3 minutes.

Neutral Task

Prompt: *"Please describe your daily routine yesterday."*

Time: 2–3 minutes.

The design also incorporates several features to ensure internal validity. First, both environments are iden-

[25]These variables were chosen based on their theoretical and empirical relevance to the development of stress-adaptive skills, ensuring that treatment and control groups are balanced across critical dimensions of adversity and economic background. To measure adversity, we will use The Adverse Childhood Experience (ACE) Questionnaire for Adults, which captures exposure to adverse experiences such as neglect, abuse, or household dysfunction during childhood.

[26]Sociological theory highlights that self-reflection enables individuals to monitor and adjust their identity and behavior across social contexts (Falk and Miller, 1998). Empirical work shows that structured self-reflection (e.g. in career courses) can increase adaptability outcomes (Kim et al., 2020).

tical in task content and timing; the only difference lies in the external conditions. Second, effort is incentivized through performance-based lotteries tied to each participant's own performance rather than to others, preventing competitive dynamics from influencing motivation. To further decouple incentives from monetary considerations, lottery entries translate into tickets for non-monetary prizes. Third, to ensure that instructions are understood and that participants remain engaged across the protocol, we include brief comprehension and attention checks. Finally, a short post-survey measures perceived stress, perceived noise, and subjective task difficulty, which allows us to validate the salience of the environmental manipulation and to better understand how participants experienced the testing conditions.

IX. E.1 Power Calculations

The experiment will be conducted at the REBEL behavioral lab at Universidad del Rosario.[27] The study will recruit 800-100 young adults out of the lab database, of approximately 20 years old from all socio-economic backgrounds in Colombia.[28]. Participants will be compensated with $15 for their participation.

IX. E.2 Stratification

In this project, randomization will be stratified based on socioeconomic status and childhood adversity, two variables linked to heterogeneity in stress responses and skill expression. This approach is motivated by the need to ensure balance on critical observables in a relatively small sample for the pilot, a setting where simple randomization may lead to imbalances that reduce power or complicate interpretation(Bruhn and McKenzie,2009);Glennerster and Takavarasha, 2013)

At baseline, participants will complete a demographic survey and the Adverse Childhood Experiences (ACE) Questionnaire for Adults, a validated instrument that captures exposure to childhood adversity, including neglect, abuse, or household dysfunction (see Figure **??**). Participants then will be stratified into four

[27]The lab has two physical locations: allowing for simultaneous experimental sessions. The facility includes a monitoring area, and interview rooms equipped with biometric measurement tools such as eye-tracking, facial recognition readers, and GSR technology. For more information: https://urosario.edu.co/en/rebel

[28]We estimated minimum detectable effects (MDEs) for a within-subject (paired) design using $\alpha = 0.05$, power of 80%, and outcomes standardized to have a variance of 1. Because each participant completes both the Structured and Adverse conditions, the power calculation is based on a paired-means test rather than a two-sample comparison. Allowing for a range of plausible within-person correlations ($\rho \in 0.3, 0.5, 0.7$), the required sample sizes to detect MDEs of 0.11 SD range from 393 - 911 participants. For plausible effect sizes, we rely on Wantchekon and Zhang (2024) who provide evidence that adverse economic conditions, such as weather shocks, promote the development of noncognitive skills like self-efficacy, self-esteem, and an internal locus of control. Their findings suggest that the effect sizes can range from 0.10 to 0.40 standard deviations, depending on the specific outcome measured

groups based on combinations of SES[29] (above/below the median) and ACE scores[30] (low/high). Random assignment to treatment conditions (structured vs. adverse environments) will occur within each stratum, ensuring that the treatment and control groups are balanced across these critical dimensions (see Figures **??** and 15).

These variables were selected given their relevance to the study's hypothesis that early adversity can shape stress-adaptive skills. Stratifying on both SES and ACE ensures that the experimental design can detect differential treatment effects across adversity profiles without confounding these with baseline economic status.

IX. E.3 Anticipated results

Based on our initial findings and the literature, it is expected that under standard conditions, students from low-adversity backgrounds will outperform their high-adversity peers on cognitive tests. However, in adverse environments, high-adversity students may exhibit adaptive skills that narrow the performance gap or even outperform low-adversity students. When the gap in performance between high and low adversity groups decreases, it can occur due to one of three reasons: (1) high-adversity children improve their performance, narrowing the gap relative to their low-adversity peers; (2) low-adversity children experience a performance decline, reducing the disparity; or (3) both groups converge due to simultaneous changes in opposite directions. These cases reflect distinct mechanisms that influence performance dynamics under different environmental conditions, highlighting how adversity may shape resilience or amplify vulnerability.

IX. E.4 Instruments

Building upon insights from prior psychological literature, this experiment aims to assess both cognitive and stress-adapted skills: attention shifting (Blair and Raver, 2014; Evans et al., 2013; Mittal et al., 2015), working memory updating (Young et al., 2018), resilience, and creativity (Dahlman et al., 2013; Fry, 2018). In addition, this study will include a measure of adaptive intelligence (Sternberg, 2021). I will use existing instruments to capture these skills.

[29] Socioeconomic Status (SES): Participants will be grouped based on an index of household socioeconomic status, constructed from detailed questions on housing quality, access to basic services, and asset ownership. This index serves as a proxy for long-run economic conditions and allows us to distinguish between economic disadvantage and other forms of adversity.

[30] ACE scores will be used to classify participants into low- and high-adversity categories. This measure captures non-economic sources of early-life stress, helping to isolate the specific contribution of adversity—distinct from SES—to cognitive and socioemotional performance under environmental stress.

Executive Function (EF): Attention shifting and Working memory updating

Executive Function refers to skills that allow individuals to plan, focus attention, remember instructions, and juggle multiple tasks (Center on the Developing Child, 2012). Cognitive science research has shown stronger cognitive function is correlated with better job market outcomes, and success in school (Bailey, 2007; Borella et al., 2010; Duncan et al., 2007). I will collect data on three types of Executive Functions: working memory, cognitive flexibility, and inhibitory control.

1. Cognitive flexibility: To measure the cognitive flexibility component from EF I will use the Wisconsin Card Sorting Test (WCST), which involves flexibly switching between different tasks or mental sets. This will allow me to measure the **attention-shifting** ability of the participants. People who are good at shifting are better at allowing their responses to be guided by the current situation rather than by an internal goal (Mittal et al., 2015).

2. Working memory: I will assess **working memory updating** using the Digits Forward and Backward subtests (WAIS-R or WAIS-III). This is a fundamental component of executive function, enabling individuals to actively maintain and manipulate information to achieve their goals and navigate their environment effectively.

3. Inhibitory control: is the deliberate overriding of dominant responses, to measure this I will use Stroop Test.

Task	**Domain**	**Description**
Digits Forward and Backward subtests (WAIS-R or WAIS-III)	Working memory	In the Digits Forward subtest, the participant must repeat the numbers dictated by the examiner in the same order. In the Digits Backward subtest, the patient must repeat the same numbers in reverse order.
Stroop Test	Inhibitory Control	The test consists of three conditions: in the first condition, the subject must say, as quickly as possible, the names of the colors that are arranged on a card. In the second condition, the subject must say the names of the colors that the words "all," "today," etc. are printed in. In the third condition, the participant has to name the colors that the words "yellow", "red", etc. are printed in.
Wisconsin Card Sorting Test (WCST)	Cognitive Flexibility	The test consists of cards that need to be classified according to color, shape or number categories. The categorization rules change every time 10 (out of a maximum of 128) response cards have been sorted correctly.

Source: based on (Faria et al., 2015)

Standard Cognitive test

To assess cognitive performance, I will rely on the Raven's matrices test, which measures individuals' capacity to think logically and solve problems in novel situations, independent of acquired knowledge. This task is a nearly universally accepted measure of fluid intelligence and a common component of IQ tests (Raven, 1936).

Resilience

Resilience will be measured using a Behavioral Persistence Task titled "Glitched Anagrams," used in Schilbach (2019)[31]. This instrument specifically captures behavioral recovery following an unexpected environmental "shock"—an intentional system error that deletes current progress, forcing the participant to restart a difficult anagram. Unlike self-reported measures of resilience, this task provides a metric of resilience based on behavior.

By introducing this "wasted effort" shock alongside background noise and input lags, the task isolates the participant's capacity to bypass frustration and maintain cognitive throughput in the face of sunk-cost disruptions. This behavioral measure directly addresses the "Hidden Strength" hypothesis by testing if students from high-adversity backgrounds exhibit a lower drop-off in effort when the system becomes unreliable.

Creativity

I will use the Alternative Uses Task (AUT) developed by Guilford in 1967 and later incorporated in the Torrance Tests of Creative Thinking (TTCT) (Torrance, 1968, 1998)[32]. The goal is to generate as many uses as possible for familiar objects in a minute. One advantage of the unusual uses task for our purposes is that it provides a clean numerical measure of creative output (Bradler et al., 2019)).

The test can be administered in different formats. However, I will follow the procedure implemented by Bradler et al. (2019) and Ashraf et al. (2020) in an experimental context. In my setting, I will ask students to identify unusual uses for three different items: (1) paper, (2) tin can, and (3) cord. In the present study, the child was scored on the number of acceptable uses he could come up with. The answers will be evaluated using the three standard measures of the unusual uses task: validity, flexibility, and originality (Guilford, 1967).

[31] Schilbach (2019) uses Anagrams and Counting Tasks as Real-Effort Persistence Measures to track self-control and willpower under the influence of stressors (like alcohol or sleep deprivation).

[32] This is the most widely used tool designed by psychologists to measure student creativity. I could also include the entire TTCT instrument, but for time restrictions, I limit it to the AUT task that has been previously used in experimental settings.

Adaptive Intelligence

Adaptive intelligence is the intelligence one needs to adapt to current problems and anticipate future problems of real-world environments (Sternberg, 2021). I hypothesize that this is a mediator to be able to use the adaptive skills regardless of the environment. To measure I will follow the abbreviated test of adaptive intelligence developed and validated by Sternberg (2021), where I present participants with a serious and real problem in the real world.In each case, respondents were asked the following questions:

- What could you do personally about the problem?
- What is wrong with what is being done (or not done) about it now?
- What would you recommend those in authority do about it?
- What are the obstacles to doing what needs to be done about it?