Juan Andrés Russy

✉ jrussycervera@gmail.com • in juan-andres-russy

Education

Ph.D. in Economics **2026**
Massachusetts Institute of Technology (MIT)

M.A in Economics **2024**
Universidad del Rosario
Excellence Diploma. GPA: 4.8/5 - USA scale: 4.0 (WES Course by Course Evaluation)

B.A in Economics **2023**
Universidad del Rosario
Minor in Applied Mathematics and Computer Science. GPA: 4.7/5 - USA scale: 4.0 (WES Course by Course Evaluation)

Working Papers & Work in Progress

The Efficiency Frontier of Public Transit
with Milena Almagro, Juan Camilo Castillo, Nicola Rosaia and Tobias Salz

Simplified Tax Regimes: A Doorway to Tax Evasion
with Monica Calijuri, Andrea Lopez-Luzuriaga and Carola Pessino

Influence of Entertainment Media on Beliefs
with Pinar Yildirim, Greg Martin and Ruben Durante

Online Communities and Margin of Contribution
with Pinar Yildirim, Martina Pocchiari and Camilo Garcia Jimeno

Experience

University of Pennsylvania
Predoctoral Fellow for Prof. Pinar Yildirim and Juan Camilo Castillo *Jul 2024 – July 2026*

Work on research exploring optimal public transportation policies and examining how media, technology, and automation influence labor markets and broader societal outcomes.

· **The Efficiency Frontier of Public Transit** *Prof. Castillo et al.*
I implemented a public transit supply-side model to study the untapped efficiency potential of public transportation, mathematically defined by the trade-off between operational cost per passenger and passenger waiting time. My contributions enable a granular analysis at the street-network level, allowing for the derivation of precise movement patterns for transit vehicles. I helped scale the optimization to any city in the US while keeping the theory tractable, deriving simple analytical formulas that elucidate the forces behind the observed patterns. Additionally, I proposed and calculated an "inefficiency index" to quantify performance between and within cities.

· **Influence of Entertainment Media on Beliefs** *Prof. Yildirim et al.*
I theoretically expanded the Gentzkow, Shapiro, and Taddy (2019) model of congressional language partisanship from a binary classification to a continuous measure of political alignment. I applied this novel methodology to analyze the partisanship of the language in entertainment shows and news transcripts via LexisNexis. Concurrently, I constructed social belief indices at the county level using 68 subthemes from Gallup Survey data, standardizing questions to follow clear ideological directions (e.g., "more support for abortion"). Finally, I implemented an IV specification exploiting random weather fluctuations on show release days—which exogenously drive home viewership—to identify the causal effect of streaming content on these specific beliefs.

· **Effects of Online Dating Platforms on Marriage and Health Outcomes** *Prof. Yildirim et al.*
I engineered a comprehensive panel dataset of dating app usage from 2002–2024 by harmonizing over 10 TB of desktop and mobile tracking data from Comscore, Tapestri, and Dewey. During this process, I identified and resolved critical data quality issues related to inconsistencies in market share data within the Tapestri dataset. I utilized this panel to identify the causal effect of online dating on marriage, divorce, and STD outcomes, analyzing heterogeneous effects based on age and the penetration of relationship-minded versus

LGBTQ+ apps. To ensure rigorous identification, I implemented advanced weak instrument robustness checks, including Anderson-Rubin confidence sets, the CLR test, and the Kleibergen-Paap Wald statistic.

- **Career Values for Labor Markets: Evidence from Robot Adoption** *Prof. Yildirim et al.*
  I processed 17 million U.S. resumes to estimate heterogeneous probability transition matrices across all occupations between 2000 and 2016 by gender, education, and experience level. I constructed heterogeneous career values—defined as the present value of a career—to measure the impact of robot automation shocks. Furthermore, I constructed individual-level trajectories of promotion and demotion, which I used to run regressions analyzing the effects of robot adoption on career paths.

- **Online Communities and Margin of Contribution** *Prof. Yildirim et al.*
  I reconstructed the conversation structure of Reddit threads across over 700 subreddits from 2010 to 2018. I investigated the relationship between post characteristics (e.g., grammar, sentiment) and thread structure (e.g., size, depth). To analyze the causal effect of corporate scandals on community discourse, I curated a dataset of 1,729 scandals and matched affected firms with their main competitors. I examined the similarity in language and sentiment following a scandal by generating and analyzing 3 TB of vector embeddings for all posts and comments using OpenAI's embedding models.

- **Optimal Urban Transportation Policy: Evidence from Chicago** *Prof. Castillo et al.*
  I developed the comprehensive replication package for the project's revision at *Econometrica*. This involved unifying previously disorganized code flows into a streamlined pipeline reproducible across R, Python, Stata, and Julia. Additionally, I generated synthetic waiting-time data to preserve the statistical properties of the original dataset while complying with strict NDA protections on Uber data.

- **Persuasion and Dissuasion in Political Campaigns: Communication and Media Coverage in Senate Races** *Prof. Yildirim et al.*
  I reviewed a non-cooperative game theory model regarding political persuasion and the dissuasion of media coverage. My contributions focused on clarifying the exposition for a revision at *Quantitative Marketing and Economics*.

Stanford University
Research Assistant, Environmental Change and Human Outcomes lab *Sep 2023 – Jun 2024*

- **Revisiting Africa's Growth Narrative with High-Resolution Microspatial Data** *R. S. Brandon de la Cuesta et al.*
  Developed and implemented methods to correct bias caused by using predictions of remotely sensed variables in inference. I identified a gap in bias-correction methods when applied to political economy settings and developed aggregation-consistent corrections to measure wealth and growth outcomes.

Universidad del Rosario
Research Assistant for Prof. Andrea Lopez-Luzuriaga *Feb 2023 – Jul 2024*

- **Simplified Tax Regimes: A Doorway to Tax Evasion** *Prof. Lopez-Luzuriaga et al.*
  I adapted and estimated a structural bunching model to disentangle the simultaneous decision between evasion and simplicity near the revenue threshold. I use these estimates to satisfy the required identification assumptions of my proposed matching strategy. I identify the ATT of Brazil's simplified regime on revenue underreporting using the universe of firms in Brazil that belong to the simplified tax regimes.

Universidad del Rosario
Research Assistant for Prof. Juan F. Vargas *Jun – Aug 2023*

- **A Peace Baby Boom? Evidence from Colombia's Peace Agreement** *Prof. Vargas et al.*
  Executed the empirical revision for a R&R decision at the *Journal of Human Resources*. My contributions included extending the temporal scope of the analysis by integrating recent Vital Statistics and conflict data, validating the integrity of the original code base, and conducting comprehensive robustness checks.

Universidad del Rosario
Research Assistant for Prof. Santiago Saavedra *Jun 2021 – Jan 2023*

- **Technology and State Capacity: Experimental Evidence from Illegal Mining in Colombia** *Prof. Saavedra*

I constructed an automated pipeline to process NICFI satellite imagery using Google Cloud and Earth Engine. I fine-tuned a ResNet deep learning model to detect open-pit mining, achieving a False Negative Rate of over 90%. on human-labeled data. I deployed this model to generate and publish monthly predictions for the Colombian Mining Monitoring Initiative. I constructed a high-resolution spatiotemporal panel (300k+ $128m \times 128m$ grids over 7 years) to enable robust inference for the project, currently under R&R at *AEJ:Applied Economics*.

Accenture Song
Data Scientist Intern *Jul 2022 – Jan 2023*

· **ML Sentiment Analysis for Colombian Slang**

Orchestrated a project to classify sentiment in Colombian slang. To address the scarcity of labeled data, I designed and built a mobile web application that incentivized internal employees to generate training labels in exchange for rewards. I utilized this dataset to train and benchmark multiple architectures, including fine-tuning a pre-trained BERT model (BETO) and developing LSTM models from scratch. Finally, I deployed the optimal model as a serverless API using AWS Lambda for internal company analytics.

Universidad del Rosario
Research Assistant for Prof. Andrea Lopez-Luzuriaga *Sep 2021 – Aug 2022*

· **When the taxman shutters the store next door** *Prof. Lopez-Luzuriaga*

This study estimates the impact of nearby business closures on family-owned firms in Ecuador. To identify family structures, I co-developed namegraph, a Python library that processes Spanish surnames and constructs genealogical networks. I implemented Breadth-First Search (BFS) algorithms (up to 7 degrees of consanguinity) and centrality measures to map a connected graph of 16 million individuals. To identify nearby closures, I developed a distributed scraping architecture utilizing 50 physical nodes, which were coordinated via AWS (SQS and DynamoDB). This system enabled us to geolocate over 1 million unformatted addresses at nearly zero cost.

Awards and Scholarships

MIT Departmental Fellowship
Full Tuition and Stipend *2026*

Graduate Scholarship, Universidad del Rosario
25 % Tuition Discount on Master's in Economics *2023*

Top 0.1 % Pruebas Saber Pro, Ministry of Education, Colombia
Colombian undergraduate standardized test *2023*

Best GPA Scholarship, Universidad del Rosario
100 % Tuition on Economics Undergraduate *2020*

Academic Excellence Scholarship, Universidad del Rosario
50 % Tuition on Economics Undergraduate *2019, 2021, 2022*

Superate con el Saber Matematicas
Finalist of Colombian National Math Olimpiads *2012*

Skills

Language: Spanish (Native), English (Fluent)

Coding: Julia, Python, R, Stata, SQL, LaTeX, Matlab, C++

Tools: GitHub, AWS, GCP, Redivis, Earth Engine, SLURM