

Early Warning Systems 2.0: Machine Learning Models to Predict and Prevent Violent Outbreaks

Daniel Mejía

Juan Diego Heredia

Economics Department, Universidad de los Andes

Organized Crime and Development in Latin America and the Caribbean - Kick-off discussion seminar

Interamerican Development Bank, Oct. 22 and 23, 2025

Motivation: *Defensoria del Pueblo*'s SAT

- The **Early Warning System (SAT** by its Spanish acronyms) **is a qualitative monitoring and warning tool aimed at identifying risk situations associated with human rights violations.** This System is administered and operated by *Defensoria del Pueblo de Colombia* (Colombian Ombudsman's Office).
- These risk situations can be understood as atypical violence events or spikes in violence in certain territories (municipalities) and periods.
- Created in 2001, the SAT has operated primarily through qualitative methods and field-based intelligence gathering, where skilled analysts deployed to vulnerable regions, monitor risk indicators, and produce qualitative risk assessments to alert authorities of impending threats to human rights.
- The SAT has enabled *Defensoria del Pueblo* to issue timely warnings that have saved lives and guided preventative actions.

Motivation (cont.)

- However, as security challenges become more complex and data availability grows, *Defensoria del Pueblo* has recognized the need to augment their proven qualitative approach with advanced analytical tools that can further improve the precision and foresight of their Early Warnings System.
- **The objective of our project is to complement the (mainly qualitative) SAT with machine learning tools aimed at forecasting atypical violent spikes.**
- By leveraging machine learning models to analyze large volumes of administrative, geospatial, and historical data, our project will strengthen and complement the SAT by enabling *Defensoria del Pueblo* to anticipate atypical surges in violent crime with greater accuracy and advance notice.

Objective(s) of our project

- **Primary objective: Develop and validate data-driven early warning machine learning predictive models of atypical violent events.**
- In practical terms, integrating these predictive models into *Defensoria del Pueblo's* workflow will enhance the precision of their alerts and allow them to identify emerging violent trends earlier than is currently possible.
- In short, our project will modernize and strengthen the SAT, combining the best of data-driven modeling with their existing qualitative expertise to improve both the timing and targeting of violence prevention efforts.

Expected results

- The project will produce risk scores, alerts, and maps that the *Defensoría's* SAT can integrate into its regular workflow, enabling earlier and more targeted activation of preventative measures.
- Comparing the model's predictions with SAT's historical alerts will help identify blind spots in coverage and refine institutional approaches and analytical attention allocation.
- If effective in Colombia, this data-driven early warning approach can serve as an innovative model for other countries in the region facing similar organized crime and violence dynamics.

Innovations

- While machine learning models have the potential to successfully forecast violent incidents in Colombia (Bazzi et al., 2022), no prior effort has been integrated into an official early warning system like the SAT.
- Our project offers a methodological innovation by translating multi-source data into actionable risk alerts, directly addressing subnational organized crime monitoring, predictive tools for risk and protective factors, and evaluating policy responses.

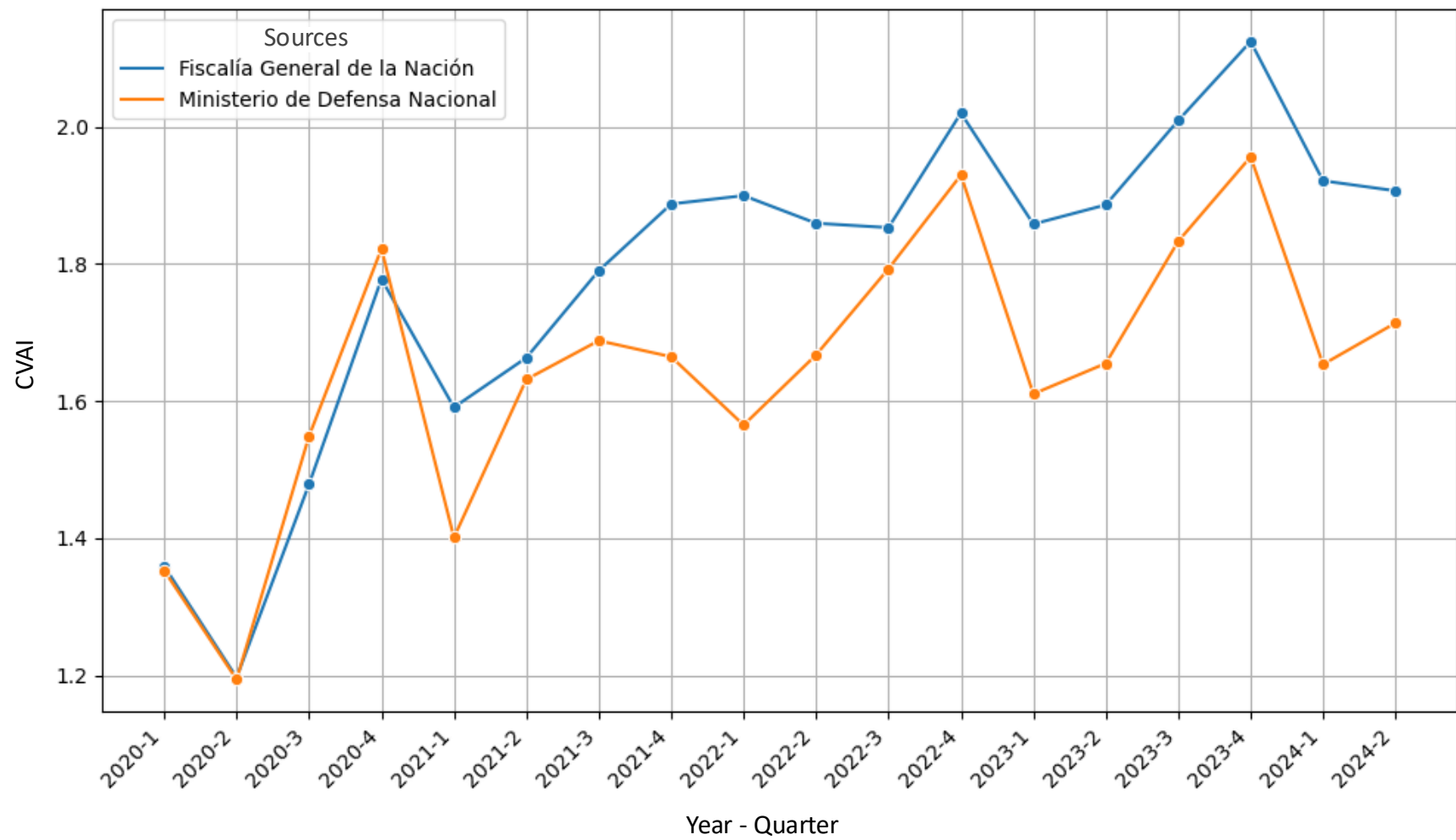
Data: Main aggregate violence indices

- **Crime and Violence Aggregate Index:** A composite measure that summarizes the intensity of violence at the municipal level per quarter. It aggregates five types of violent crimes commonly associated with organized crime, weighted by their social impact (based on legal penalties in the Colombian Penal Code): homicides (17.04%), extortion (10.31%), kidnapping (14.35%), terrorism (13.45%), and massacres (44.84%). The index is normalized by population to enable cross-municipal comparisons.
- **Intimidation Index:** actions by organized criminal groups aimed at creating fear among the population. It incorporates threats, attempted murders and attacks, forced displacement, and harassments. This index captures the extent to which violence is used to intimidate local populations.
- **Criminal Governance Index:** measures territorial control exercised by criminal organizations. It includes forced confinements, illegal checkpoints, armed strikes, and extortion. This index reflects the degree to which criminal actors impose governance structures and restrict civilian mobility.

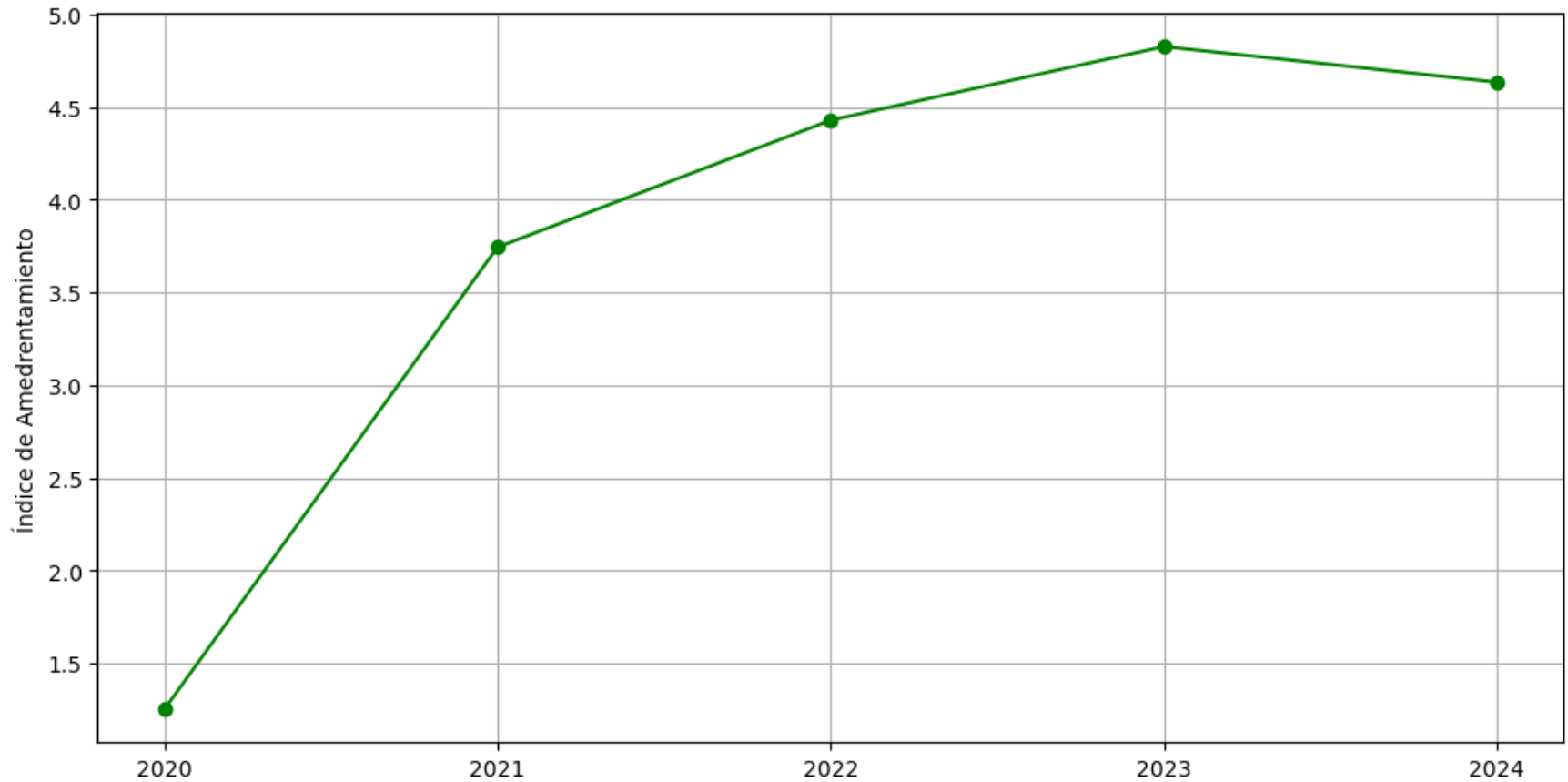
Descriptive statistics

	Number of Observations	Average	Standard Deviation	Minimum	Maximum	Proportion of atypical events at the municipal- quarter level: median+1 s.d. (%)	Proportion of atypical events at the municipal- quarter level: median+2 s.d. (%)
Terrorism	2,687	2.2	2.1	1.0	29.0	17.4	8.75
Extortion	24,890	4.7	18.3	1.0	729.0	21.88	11.84
Homicides	47,997	6.7	24.8	1.0	614.0	17.98	8.53
Massacres	259	3.8	1.9	3.0	18.0	0.0	0.0
Kidnappings	10,538	2.7	4.6	1.0	179.0	14.46	8.05
CVAI	67,273	1.4	2.2	0.0	53.5	17.05	8.62
IGC	26,920	0.0	0.3	0.0	10.0		
IA	26,920	0.2	1.1	0.0	46.0		

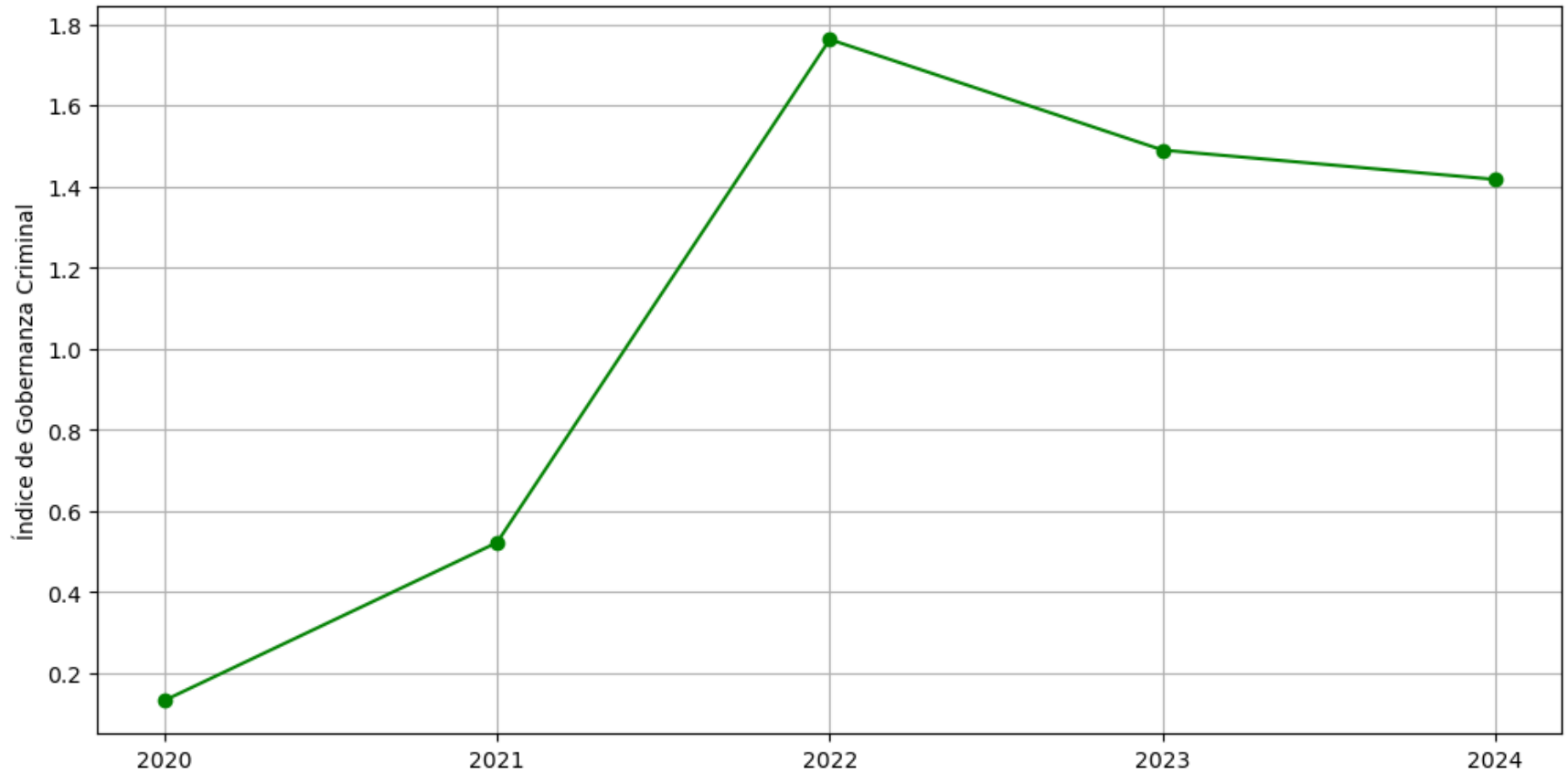
Aggregate Crime and Violence Index, 2020I-2024II



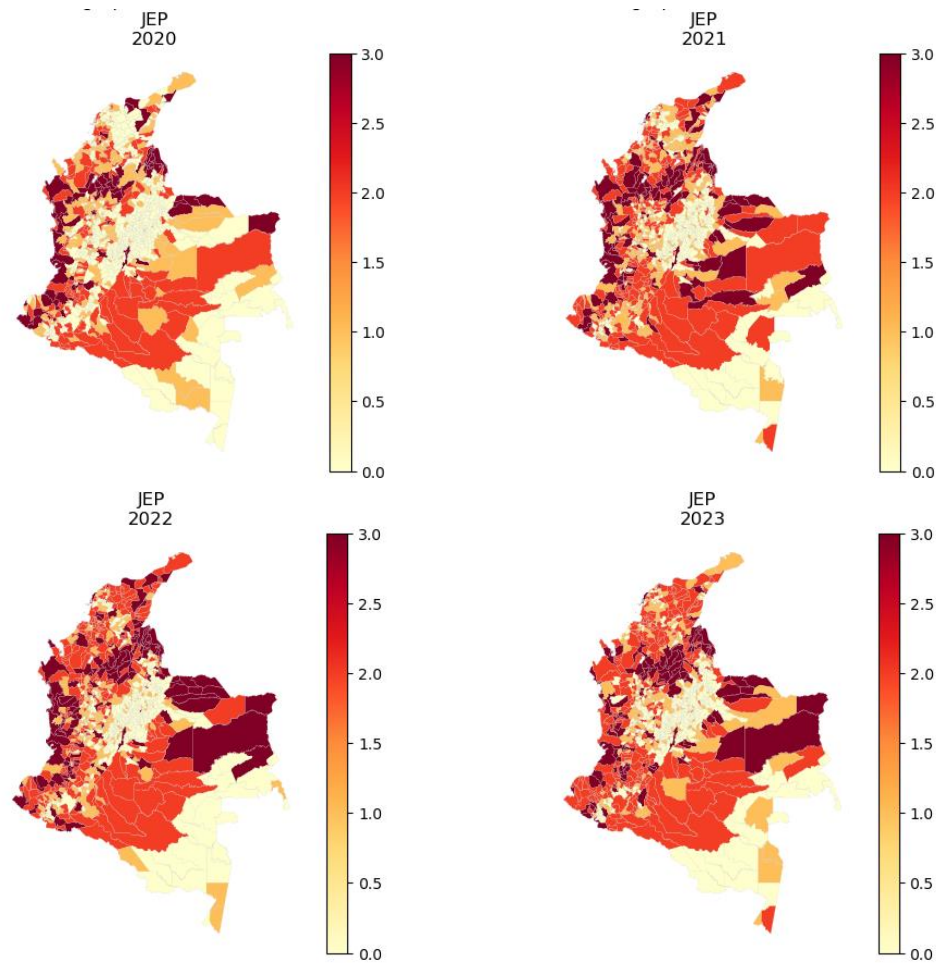
Intimidation Index, 2020-2024



Criminal governance index, 2020-2024



Number of criminal groups per municipality in Colombia: 2020 - 2023



Methodology

Specific objective:

- Predict the probability of atypical violent events occurring in a given municipality in a given quarter.

Why classification and not regression?

- The focus is on a binary variable: atypical violence (yes/no).
- Classification allows estimating probabilities associated with risk.
- Models can be fine-tuned to prioritize sensitivity or precision.
- Easy to integrate into *Defensoria's* SAT.

Machine learning models

- **Elastic Net:** A linear regression approach that combines L1 (Lasso) and L2 (Ridge) penalties to automatically select relevant variables while reducing overfitting.
- **Random Forest:** An ensemble method that builds multiple decision trees to handle high-dimensional data and reduce overfitting through averaging predictions.
- **XGBoost:** A boosting algorithm with iterative optimization and regularization that efficiently handles imbalanced datasets and missing values, similar to Random Forest but with enhanced performance.
- **LSTM Networks:** Recurrent neural networks designed to model temporal dynamics more effectively than traditional methods by detecting critical patterns in time-series data.

Model evaluation

Confusion Matrix

Confusion Matrix	# of non-atypical violent events observed	# of atypical violent events observed
# of non-atypical violent events predicted	True Negative ✓	False Negative ✗
# of atypical violent events	False Positive ✗	True Positive ✓

➤ Evaluation metrics:

- **Accuracy:** “How well it predicts violence in general”
- **Sensitivity (Recall):** “Not missing any critical event”
- **Precision:** “Not allocating preventive resources unnecessarily”
- **AUC:** “Ability to prioritize correctly”

Data sources

- Data on violence in Colombia comes from three main sources:
 - **Attorney General's Office:** Monthly records of homicides, extortion, kidnappings, terrorism, and massacres at the municipal and monthly level.
 - **Ministry of Defense:** Historical series of the same crimes with broader temporal coverage, at the municipal and monthly level.
 - **Special Jurisdiction for Peace (JEP):** Data on armed groups' presence and violent events, including forced displacements, harassment, and armed strikes and many other actions carried out by organized criminal groups.
 - **Presence of organized criminal groups at the municipality level (JEP, FIP).**
- Socioeconomic indicators at the municipality level in Colombia:
 - **CEDE Municipal Panel:** Demographic, socioeconomic, and institutional variables, such as poverty indices, inequality, institutional capacity (data at the year level).
 - **Illicit coca crops and gold mining:** Main sources of funding for illegal armed groups (data at the year level).
 - **Nighttime lights (VIIRS Nighttime Light):** Proxy indicator of local economic activity.

Next steps

- Update all data sources through mid-2025 and build the panel with observations at the municipality – quarter (year) level.
- Initial estimation of the main machine learning models and evaluation.
- Compare the main predictions from the different models' estimations with *Defensoria's* SAT.
- We will be working closely with *Defensoria's* SAT team throughout the project.