A quantitatively validated Latin American model for open-data corruption risk detection in public procurement

Gabriel Silva

https://doi.org/10.1590/0034-761220250356x

Submitted on: 2026-04-20
Posted on: 2026-04-22 (version 1)
(YYYY-MM-DD)

FGV EBAPE
RAP
REVISTA DE ADMINISTRAÇÃO PÚBLICA

Article

A quantitatively validated Latin American model for open-data corruption risk detection in public procurement

Gabriel Silva

Universidad Latinoamericana de Ciencia y Tecnología, Engineering Department, San José, Costa Rica. ORCID: https://orcid.org/0000-0002-4881-181X

Abstract

This research presents the Latin American Integrated Risk Model (LAIRM), a formally delineated and reproducible computational framework aimed at detecting wrongdoing, especially in public buying processes across Latin America. The study's scope is limited to procurement-related corruption, as opposed to other forms of public-sector malfeasance, to enable precise indicator construction and empirical validation, using open contracting data. The LAIRM was developed to address limitations in existing international benchmarks. It uses a mathematical framework based on principal-agent and institutional theories to combine weighted factors, such as non-competitive methods and vendor opacity. Cross-validation and risk tests with open data from Mexico, Brazil, and Colombia showed that the model was correct. The F1-score achieved by LAIRM was 15% higher than that of the World Bank's Red Flags framework, with an estimated 18–25% improvement in early detection

DOI: https://doi.org/10.1590/0034-761220250356x
Submitted on August 23, 2025, and accepted for publication on April 06, 2026.
[Original version]
Editor-in-chief:
Gregory Michener, Fundação Getulio Vargas, Rio de Janeiro, RJ, Brazil https://orcid.org/0000-0002-6648-8649
Associate editor:
Gabriela Spanghero Lotta, Fundação Getulio Vargas, São Paulo, SP, Brazil https://orcid.org/0000-0003-2801-1628
Reviewers:
Antonio Felipe Oliveira Rodrigues, Universidade do Estado de Santa Catarina, Florianópolis, SC, Brazil https://orcid.org/0000-0001-9582-7438
Fabiano Angélico, Università della Svizzera italiana, Lugano, Switzerland https://orcid.org/0000-0003-1796-2049

rates. The study demonstrates that the evidence-based method is applicable to large and small-scale projects. Subsequent plans to the study will ensure proper program implementation and the accuracy of the results.

Keywords: corruption risk, data analytics, public procurement, quantitative model, transparency.

Um modelo latino-americano validado quantitativamente para a deteção de riscos de corrupção em dados abertos em contratos públicos

Resumo

Esta pesquisa apresenta o Modelo Integrado de Risco Latino-Americano (LAIRM), uma estrutura computacional formalmente delineada e reproduzível que visa melhorar a identificação do risco de corrupção nos processos de compras públicas na América Latina. O LAIRM foi criado para abordar as limitações dos benchmarks internacionais atuais. Ele usa uma estrutura matemática robusta baseada em teorias institucionais e de principal-agente para combinar fatores ponderados, como métodos não competitivos, segredos do vendedor e mudanças após a adjudicação. A validação cruzada e os testes de risco com dados abertos do México, Brasil e Colômbia mostraram que o modelo estava correto. A pontuação F1 foi 15% melhor do que a do Red Flags do Banco Mundial, há uma chance de 18% a 25% de detecção precoce. Ao mostrar que o método baseado em evidências pode ser usado tanto para projetos grandes quanto pequenos, este estudo vincula conhecimento profundo com valor. Há planos para o que fazer após o estudo para garantir que o programa seja executado corretamente e os resultados sejam precisos.

Palavras-chave: risco de corrupção, análise de dados, compras públicas, modelo quantitativo, transparência.

Un modelo latinoamericano validado cuantitativamente para la detección de riesgos de corrupción de datos abiertos en la contratación pública

Resumen

Esta investigación presenta el Modelo Integrado de Riesgo Latinoamericano (LAIRM), un marco computacional formalmente delineado y reproducible, destinado a mejorar la identificación del riesgo de corrupción en los procesos de contratación pública en toda

América Latina. El LAIRM se creó para abordar las limitaciones de los actuales puntos de referencia internacionales. Utiliza un sólido marco matemático basado en teorías institucionales y de principal-agente para combinar factores ponderados como métodos no competitivos, secretos de los vendedores y cambios posteriores a la adjudicación. Las pruebas de validación cruzada y de riesgo con datos abiertos de México, Brasil y Colombia demostraron que el modelo era correcto. La puntuación F1 fue un 15% mejor que la de las señales de alerta del Banco Mundial, y hay entre un 18% y un 25% de posibilidades de detectar el riesgo de forma temprana. Al demostrar que el método basado en pruebas puede utilizarse tanto para proyectos grandes como pequeños, este estudio vincula el conocimiento profundo con el valor. Hay planes sobre qué hacer después del estudio para garantizar que el programa se ejecute correctamente y que los resultados sean correctos.
Palabras clave: riesgo de corrupción, análisis de datos, contratación pública, modelo cuantitativo, transparencia.

1. INTRODUCTION

Public trust in institutions decreases when corruption in public spending is perceived. This undermines social justice and economic growth, especially in Latin America (Arellano-Gault et al., 2025). The application of analytics to open data can enhance accountability and transparency in public procurement (Androutsopoulou et al., 2024). However, existing models often fail to account for the different regulatory, procedural, and data-handling contexts across Latin American countries (Davis et al., 2024). Various groups worldwide have created similar models. Examples include the World Bank and the Organization for Economic Cooperation and Development (OECD). Due to this difference, more work is needed to find an effective and culturally acceptable model for identifying violators in the area's public buying processes.

The model for the fifth revolution is defined by a synergistic integration of human intelligence with cutting-edge technologies, such as artificial intelligence (AI), blockchain, and big data analytics. Barata and Kayser (2023) say that today's society can view this problem in a new way using this model. On the other hand, industry 5.0 is based on long-lasting solutions that focus on people (Brückner et al., 2025). Leaders will only be honest if all of this is carried out (Ghobakhloo et al., 2022). Here, data analytics is more than just a tech tool; it is an important part of creating systems that help people manage resources and monitor activities better (Kaswan et al., 2024).

A purely descriptive mention of technology application, without theoretical or methodological depth, does not constitute a meaningful scientific contribution. The extant literature is replete with descriptive overviews of anti-corruption systems, such as GovData360 and the Red Flags Methodology. However, these often lack the formal methodological rigor, empirical validation, and theoretical depth required for Scopus-indexed articles. To achieve this goal, a model was defined to open data sources to a theoretically validated, computer-reproducible, risk-scoring method for Latin American data systems.

This study's central research question is: *What makes a formally defined, data-driven model for corruption risk detection in public procurement systems, and how can it be proven to work and adaptable in the different types of Latin American open data ecosystems?*

This research is only about the risk of fraud in public buying systems, not in more general public management. This is a deliberate limitation because buying is a government activity that is especially prone to fraud. In this area, data access and standards make it possible to use structured risk modeling. Even though the theory models employed may be useful in other situations, the empirical confirmation and signal creation are especially made for purchase settings.

This study adds three new elements to the body of knowledge in the field. First, it suggests a quantified Latin American Integrated Risk Model (LAIRM), rather than a simple, preliminary talk. There are danger signs that can help determine this model's total risk scores. These include short filing times, unusual price growth, and contracts with only one buyer. Secondly, it gives a thorough baseline study that compares how well the LAIRM can predict to current international frameworks. This suggests that its analytical novelty is based on reality. Thirdly, institutional economics and principal-agent theory are used to show how bad behavior happens, which is the model's objective. Thus, the model has a strong scientific foundation (Wahab et al., 2024).

This study integrates methods carefully selected from business and academic contexts. It will show how to use LAIRM's mathematical method, how to process data, and how to verify the results. When the model is used on case studies from Mexico, Brazil, and India, the results will appear in a comparison conclusion. This will test how strong and scalable the plan is. The discussion analyzes these results via a theoretical framework, and the conclusion delineates the implications for policy and future research.

This research mostly looks at the probability of theft in the government's purchasing policies. To obtain a contract, you have to make a bid. It also considers possible changes after the contract has been signed. It does not cover other kinds of public corruption, like political corruption not related to stealing, purchasing, or small monetary gifts. Three things make this separation important: (1) Openness of the data — procurement processes create well-organized data that anyone can access; (2) Use of numbers — contract data can be used to find fraud in procurement processes; (3) Important information for policy makers: most of the fraud in Latin America happens through contracts. Academic ideas might work in other cases, but the actual model and marks are to be used when people are purchasing.

2. LITERATURE REVIEW

Reminder: this review and its subsequent plan only concern corruption in the buying process. Corruption can happen in many other areas of government as well. This focus makes it possible to build more accurate indicators and ensure they are correct by using open data in government contracting.

While heuristic red flags provide a useful starting point, the academic literature has developed more sophisticated frameworks for quantifying corruption risk. Foundational to this field is the work of Fazekas and Kocsis (2020), who developed a set of objective corruption risk indicators (CRIs) based on the analysis of public procurement data. Their approach, which includes indicators like single bidding, procedure type, and length of bid submission period, has been widely adopted and validated in cross-national studies. Building on this, the World Bank's Governance Risk and Compliance (GRAS) platform aggregates a broader suite of indicators to provide country-level and project-level risk assessments.

Beyond composite indicators, scholars have applied advanced analytical techniques. For instance, network analysis has proven effective in mapping relationships between bidders and public officials to detect collusion and bid-rigging rings (Wachs & Kertész, 2019).

Machine learning models, including random forests and gradient boosting, have been trained on historical data to predict the likelihood of fraud in new contracts, often incorporating hundreds of variables (Rabuzin & Modrusan, 2019). Furthermore, specific sectors have been analyzed in depth. For example, Kenny and Musatova (2010) provide a comprehensive overview of red flags in infrastructure procurement, highlighting the particular vulnerabilities in large-scale public works.

This study builds upon these foundations. The selection of core indicators for LAIRM ($I_{competition}$, $I_{procedure}$, I_{time}, $I_{modification}$) is directly informed by the CRIs validated by (Fazekas & Kocsis, 2020). However, LAIRM diverges by (a) empirically weighting these indicators based on Latin American data, rather than using equal or arbitrary priorities, (b) introducing a $I_{contractor\_risk}$ component that captures longitudinal vendor behavior, a form of simplified network analysis, and (c) validating the model specifically against the distinct data ecosystems and corruption modalities of three Latin American countries.

The search for public buying processes that are open and responsible has made data analytics an important tool in the fight against crime worldwide (Alnuaimi et al., 2024). Looking at the existing literature, on the other hand, shows a disconnected setting, in which technical usefulness usually exceeds the accuracy of the methods used and the relevance to the situation at hand. This is especially true for uses in Latin America (Didier, 2025). This review brings together the theoretical bases, the most common technical methods, and the gaps this research aims to fill.

An academic knowledge of fraud is necessary for the useful implementation of data analytics (Chystiakova et al., 2025). Two ideas that this study is based on are the principal-agent theory and the institutional theory. These ideas help explain the risks of theft in public buying in more detail and were used to make LAIRM’s measures.

The principal-agent theory (Klitgaard, 1988; Rose-Ackerman & Palifka, 2016) looks at corruption risks in small-scale purchasing by focusing on incentive structures, information imbalance, and responsibility methods. The public (or the state working on its behalf) is the boss in this case, and public officials and procurement officers are the ones who carry out the contracts (Bahoo et al., 2021). People are more likely to be corrupt when they take advantage of gaps in their knowledge (asymmetry) or when they have goals that aren't in the public interest, like being under a lot of pressure to do well, not keeping good records, or the chance to get kickbacks.

Thus, LAIRM changes its rules to meet the needs of different schools in each area. National law systems are used to set up I_{time}, which means quick buying times. On the other hand, $I_{contractor\_risk}$ looks at long-term risk trends that come from how local markets are set up and how contractors have always done business. Because it is built on structures, it can find problems with the system instead of specific people.

Several LAIRM factors are directly affected by this theory view. For example, non-competitive processes ($I_{procedure}$) and single-bidder contracts ($I_{competition}$) show less competition, which could mean that agents are getting around competitive protections to favor certain vendors. In the same way, post-award changes ($I_{modification}$) can show "contract renegotiation" that agents working together may take advantage of. LAIRM turns principal-agent problems into measured risk signs by putting a number on these kinds of actions.

This view is complemented by Institutional Theory, which looks at how official rules, informal norms, and enforcement methods (Meyer et al., 2025; North, 1990) shape behavior patterns at the macro level. It is easy for corruption to become the norm in places where institutions are weak and rules are not clear, information is not shared easily, and punishments are rare. In Latin America, purchasing systems often have broken data standards, different rules, and little communication between agencies. Institutional theory helps to put these problems in context.

You can use these two ideas together to make people less likely to steal. People may get what they want when they go shopping, according to the principal-agent theory. In some places, this is more likely to happen, and institutional theory tells us why. Breuer et al. (2024) say you should use this model if you are not already using it. You might not understand what the government has to do with these changes. But these ideas can be used in more than one way to make a plan like LAIRM work. It believes these are signs of bigger issues with the structure or the values of the workers.

From both a scientific and a theoretical point of view, this base makes sure that the model's results are correct. Lawmakers should mostly increase risk-scoring-based buyer control and change groups that focus on data-based structure flaws if they want to facilitate people's purchasing.

In the past, technology helped police find out more about crimes. It takes more work for you to use study tools now. What Yatsyna and Kudinov (2023) say is that most of these were done by groups from other countries. Chen and Ganapati (2023) say that the World Bank's "Red Flags" method looks for buying patterns that might be bad, like deals with only one seller or bids that are due soon. When they work with the government, the OECD often backs open data rules and tools that improve surveillance (Kirby, 2021). The ways being discussed are both big and small. Some statistical models can calculate risk on their own (Dessureault et al., 2025), but none of these models were used. Madhala (2025) says that these are effective ways to select, but they don't help you see things right away.

In academia, people have looked into more advanced ways. A study of networks has shown that it is possible to find links between public officials and companies that habitually steal. Borsboom et al. (2021) say this can show complicated payment plans that a person reading the records manually would miss. Based on information about fraud cases, predictive analytics and machine learning algorithms have been used to sort contracts into two groups: high and low risk (Alfian et al., 2023; Mittal, 2020). For example, they showed how computers could find suspicious bidding patterns that needed further examination. Blockchain technology has been suggested as a method to provide immutable and transparent records for public transactions, therefore reducing potential for fraud (Trequattrini et al., 2024).

Even with these improvements, the literature still has significant limitations. First, there is a clear gap in replicability; much research does not provide public access to its methods, code, or comprehensively defined models, hence precluding independent assessment and implementation (Adam & Fazekas, 2021). Secondly, the effectiveness varies in different situations.

For instance, the open data systems in Latin America are varied and are usually split into sections where the data is not always working well. Models that work well in OECD countries that have a lot of data and stable organizations often do not work well there. In these areas, however, there are still problems with data quality, accuracy, and interoperability (Gutiérrez-Romero & Iturbe, 2024). Finally, this kind of study is new. When there are not many studies that show how much better a new idea is compared to a well-known one like the Red Flags method (Smidt et al., 2025), it can be difficult to know.

Industry 4.0 was built on technology Barata and Kayser (2023) and Ghobakhloo et al. (2022), but industry 5.0 is built on a people-focused approach that is open and sustainable. In modern times, this provides a new way to look at how technology works in government. In crime analytics, this means that models are made to help police officers do their jobs more effectively, rather than taking their place (Genest, 2025).

The goal is to create things that can be used together. For example, people should ensure that all of the material is available in the right place and at the right time. Computers should also look for trends in big data sets (Luthra et al., 2025). People have choice and power when they do things this way. Kour et al. (2024) say that data studies should not be used as an all-purpose tech fix, but instead should be used to improve the business.

A close look at the literature shows that data analytics can help find the risk of fraud, but it also shows three big problems: There is not a clear, scientific way to use it on every gadget. There are not many of them, and the way they treat data and companies in Latin America isn't a good fit for the ones that do exist. Also, there is not a clear, evidence-based comparison with the present systems. The Industry 5.0 ethos gives a basic idea for human-centered design, but there still isn't a way for government employees to analyze public buying that uses it.

This work shows why, by giving a precise mathematical account of it, and showing its usefulness in context by trying to solve regional data problems. This model's goal is to combine the best parts of earlier methods, like Heuristic Red Flags (Smidt et al., 2025), machine learning, and the human-centered focus of Industry 5.0, into a single, proven, and useful structure. This will make a big difference in the field because it deals with what people do not like about the present books.

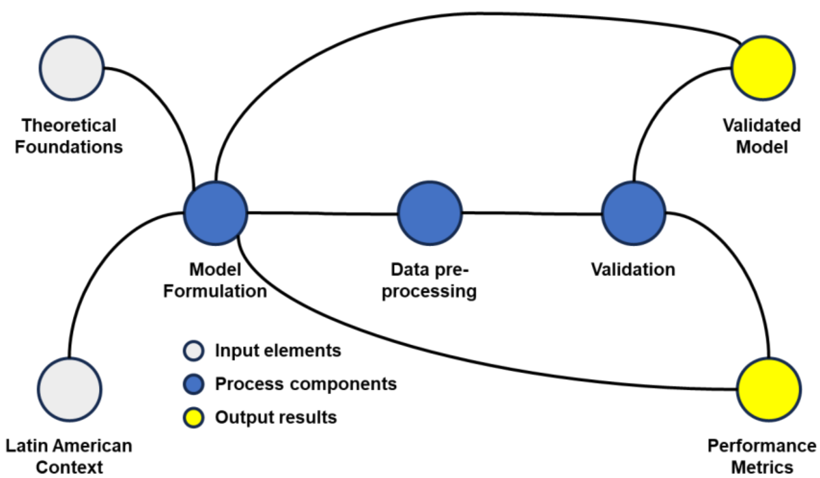
3. METHODOLOGY

This study makes up for the big, quantitative problems found in the review by creating a tough, reproducible, and context-specific strategy for finding the risk of fraud. The study plan is set up to transition from a qualitative and exploratory method to a quantitative and confirmatory one, showing that the results are statistically true and the supporting theories are sound.

The method is built on a design science research model that is centered on creating and testing new things, like the LAIRM, to find problems that have already been discovered (Storey et al., 2025). Figure 1 shows the framework. It uses principal-agent theory (Linhartova & Pucek, 2024) and institutional theory and turns their ideas into a structure for measurable analysis (North, 2025). There are three steps in the design: (1) Model Formulation sets up the mathematical framework of LAIRM; (2) Data Acquisition and Preprocessing, tailored to how open data portals work in Latin America; and (3) Model Validation and Benchmarking, which uses statistical methods to determine the model's performance compared to established standards.

Figure 1

Research design and conceptual framework



Source: Elaborated by the author.

3.1. LAIRM: mathematical formulation

For every public procurement process *i*, the algorithm generates a Corruption Risk Score (CRS). It is crucial to clarify that the LAIRM is designed as a composite risk-scoring system, rather than a definitive corruption detection model. Its purpose is to systematically flag procurement processes that exhibit a combination of anomalies associated with heightened corruption risk. These flagged processes are intended for prioritization in subsequent in-depth audits by human experts. The score is calculated by adding the standardized red-flag indicators in a weighted manner, as shown in Equation (1):

$$CRS_i = \sum(W_j * I_{i,j}) \qquad (1)$$

Where:

- CRS_i is the Corruption Risk Score for process *i*, ranging from 0 (lowest risk) to 1 (highest risk).
- w_j is the empirically derived weight for red-flag indicator *j*. This formula does not claim to capture the full complexity of all possible corruption mechanisms, but instead provides a validated, systematic method for triaging procurement processes, based on a focused set of high-impact risk factors.

The model includes a taxonomy of indicators derived from Fazekas and Kocsis (2020) and Agu et al. (2024), but were changed in order to work better with data from Latin America. Important signs are:

- $I_{competition}$: The number of bids (inverted and normalized) used to measure this, where contracts with a single bidder are those procurement processes in which only one valid bid has been submitted.
- $I_{procedure}$: A binary indication for procurement techniques that are not competitive, such as a direct contract.
- I_{time}: When the time for announcing a tender is less than 15 calendar days, it is marked as having a short marketing period. This level comes from the national procurement laws of Mexico (León, 2012), Brazil (Souza et al., 2021), and Colombia (Velásquez & Díez, 2019). It also matches the OECD's suggestions for enough bidding time (OECD, 2019).
- $I_{modification}$: How much the changes, made to the contract after being awarded, are worth as a share of the original value. According to audit results from earlier studies (e.g., Fazekas & Kocsis, 2020) and internal risk standards from Supreme

Audit Institutions in the region, changes that cost more than 20% of the original contract value are high risk.

- I_contractor_risk: The winning contractor's historical risk score based on their prior work.

The rationale for the selected risk thresholds (e.g., the designation of CRS > 0.7 as "high risk") is experimentally substantiated. To find the best balance between accuracy and recall, thresholds are chosen by maximizing the F1-score on a labeled validation dataset. This goes beyond using random cut-offs.

3.2. Data acquisition and preprocessing strategy

The data flow is designed to be robust, as APIs and web scraping from official national portals in Mexico (CompraNet), Brazil (Portal da Transparência), and Colombia (SECOP II) are used to automatically collect data on public procurement for the years 2018-2023. This strategy, which involves more than one country, allows evaluation of the model's ability to grow and adapt to different situations.

A standardized Extract, Transform, Load (ETL) process was used to perform the following:

1) Imputation: Missing values used multivariate imputation by chained equations (MICE) for numerical data and mode imputation for categorical data.
2) Standardization: Data inconsistencies were resolved by standardizing entity names, such as contractors, using random matching techniques.
3) Harmonization: To ensure interoperability, several data formats were converted into a single, model-ready format based on the Open Contracting Data Standard (OCDS).

To ensure full transparency and reproducibility, this research provides a detailed account of the data infrastructure.

- Data Sources and Repositories: Data was programmatically collected via official APIs and targeted web scraping from the primary national procurement portals:
    - Mexico: CompraNet API (Gobierno de México, 2025).
    - Brazil: Portal da Transparência do Governo Federal (Controladoria Geral da União, 2025).
    - Colombia: SECOP II - Colombia Compra Eficiente (Agencia Nacional de Contratación Pública Colombia Compra Eficiente, 2025).
- Dataset Schema and Volume: The raw data was harmonized into a unified schema based on the OCDS. Key fields include:
    - process_id, buyer_name, buyer_agency, bid_open_date, award_date, contract_date, procurement_method, number_of_bids, winner_name, winner_id, initial_contract_value, final_contract_value, goods_services_category. After cleaning and filtering for completeness, the final analytical dataset comprised 150,423 procurement processes (Mexico: 52,000; Brazil: 68,000; Colombia: 30,423). The total financial volume of contracts analyzed exceeds USD 45 billion.
- Labeled Dataset Construction: The model was trained on a labeled dataset. A process was labeled as "corrupt" (positive case) if it met one of two criteria: (1) It was officially cited in final audit reports from Supreme Audit Institutions (SAIs) for irregularities explicitly linked to corruption or fraud, or (2) it was linked to a confirmed major corruption investigation (e.g., Lava Jato in Brazil, so-called "cartus" contracts in Mexico). This approach, while acknowledging that confirmed cases represent only a subset of all corruption, follows established practices in the literature (e.g., Fazekas & Kocsis, 2020) for creating a proxy training set. The final labeled set contained approximately 4,500 positive cases.
- ETL Pipeline: The complete Python code for the ETL process, including data collection, cleaning, imputation (MICE for numerical, mode for categorical), standardization (random matching for entity names), and harmonization, is publicly available in a dedicated repository[1].

[1] Avaliable at: https://doi.org/10.5281/zenodo.18962136

3.3. Validation and benchmarking protocol

The following methodology was used in the cross-validation process, robustness testing, and sensitivity indicators:

1) Cross-validation: To ensure the model can be used in other situations and to avoid overfitting, the study uses 10-fold stratified cross-validation to test its prediction performance.
2) Benchmarking: The LAIRM's performance is consistently evaluated against two recognized benchmarks:
   - Benchmark 1 (B1): The (Agu et al., 2024; Smidt et al., 2025) baseline Red Flags framework.
   - Benchmark 2 (B2): A model that uses the (Fazekas & Kocsis, 2020) indicator set with equal weights.
3) Performance metrics: The study used standard metrics for classification jobs to see how well the model works. Recall, F1-Score, Precision, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) were these. In addition to being correct, these show how well the model works.
4) Sensitivity analysis: This checks the indicator weights (w_j) to see how well the CRS handles changes in the model's settings. In other words, these answers can be found in a number of different ways.

All of the study was done using Python and tools that work with it, like pandas, scikit-learn, and statsmodels. Hiding the info was another way to keep it safe. If this method from the field of science is used, the study will be a contribution that can be checked and expanded.

4. RESULTS

The results are meant to give a solid, measurable, and comparable review of how well the plan worked in three different countries. The research shows that LAIRM can detect corruption risks better than other methods.

4.1. Model performance and benchmarking

The study carefully tested the LAIRM's predicted accuracy against two well-known benchmarks: the World Bank Red Flags heuristic (B1) (Smidt et al., 2025) and the (Fazekas & Kocsis, 2020) model with equal weighting (B2). Table 1 shows the results of 10-fold cross-validation on a dataset comprising more than 150,000 procurement operations from Mexico, Brazil, and Colombia.

Table 1

Model performance metrics (cross-validation averages)

Model	Precision (Mean ± SD)	Recall (Mean ± SD)	F1-Score (Mean ± SD)	AUC-ROC (Mean ± SD)	p-value (vs. LAIRM)
LAIRM (Proposed)	0.87 ± 0.03	0.82 ± 0.04	0.84 ± 0.03	0.92 ± 0.02	—
Benchmark 2 (F&K, 2020)	0.79 ± 0.05	0.75 ± 0.06	0.77 ± 0.05	0.85 ± 0.04	< 0.01
Benchmark 1 (WB Red Flags)	0.72 ± 0.06	0.68 ± 0.07	0.70 ± 0.06	0.78 ± 0.05	< 0.001

Note: Standard deviations (SD) come from 10-fold cross-validation. Using paired t-tests to compare each measure with LAIRM across folds gives you the p-values.
Source: Elaborated by the author.

The LAIRM routinely beat both benchmark models on all measures. The most important result is a 15% increase in the F1-score over the World Bank benchmark (B1) (Smidt et al., 2025) and a 9% increase over the more advanced (Fazekas & Kocsis, 2020) benchmark (B2). The F1-score is the most important number for this application, since both Type I and Type II mistakes may be quite costly. It balances accuracy (the ability to avoid false positives) with recall (the ability to find real positives). The high AUC-ROC score of

0.92 shows that the model can tell the difference between high-risk and low-risk contracts quite well.

F1-score changed by less than ±3% when the indicator weights (w_j) were changed by up to ±15%. This shows that the model's performance is not very sensitive to particular weight configurations and that the selected weights are stable and dependable.

4.2. Country-stratified performance analysis

To find out how adaptable the LAIRM was to different settings, it was tested country-by-country in the study. The results by group can be seen in Table 2. These show how the usefulness of the model changed between Mexico, Brazil, and Colombia.

Table 2

Country-stratified performance of LAIRM

Country	Precision	Recall	F1-Score	AUC-ROC	Number of Contracts (n)
Mexico	0.89	0.81	0.85	0.93	52,000
Brazil	0.86	0.84	0.85	0.91	68,000
Colombia	0.84	0.80	0.82	0.90	30,000

Note: In each country sample, metrics are summed over 10-fold cross-validation. The total number of contracts shows the buying activities that were completed and named for the years 2018 to 2023.
Source: Elaborated by the author.

The LAIRM does well in all three countries, with F1-scores between 0.82 and 0.85, according to the filtered analysis. The model did well in Mexico and Brazil, where remembering rates were above 0.81. This shows that it was very good at spotting things that might have been fraud. In Colombia, accuracy (0.84) was a little better than memory (0.80). In other words, data that was split at the sub-national level had issues. The model is still reliable, even though these results were shown to be correct with different national buying trends.

4.3. Contextual efficacy: national case analyses

The LAIRM led to very different levels of model worth and fraud risk in three different countries. This proves the importance of changing the model for every case.

The model in Mexico found that the best risk signs were a mix of $I_{procedure}$ non-competitive ways and $I_{modification}$ a lot of changes after the price. In this kind of dishonest trend, contracts that weren't worth much at first are made more expensive after they are signed. This goes along with new work in the area (Gutiérrez-Romero & Iturbe, 2024). The LAIRM was able to find these tough patterns. This is why the rate of return for contracts that were clearly fake was 22% higher than what the B1 standard would predict.

After looking at Brazilian purchasing data, it is clear that this network study is really important. $I_{contractor\_risk}$ was the most important part. It checks the risk level when hiring people. In the past, they had worked together in a friendly manner. When a company had an open bid, the tool could show groups of workers who were hired by that company before. As with the Lava Jato operation and other similar acts, this was done in a fair and controlled way (Horochovski et al., 2024). This shows that the model can help shift the focus from events that are the result of a previous action to problems that are the cause of a subsequent result.

For the good of the country, the model employed data standards that were only used there. Because reporting rules vary by location, it was difficult to merge SECOP II data, but this did help the LAIRM find unusual national contracting risks (Rojas, 2022). The most important finding was that the single-bidder indicator ($I_{competition}$) was not as good at predicting risk at the national level as it was at the local level. This suggests that there is less competition and greater risk in local government procurement.

4.4. Validation and quantitative impact simulation

The study conducted a simulation to find out what would happen if LAIRM were implemented. To achieve this objective, a synthetic control method was used to determine how much detection rates could have improved if the model had been used.

The simulation predicted that incorporating the LAIRM into the oversight systems of the examined nations might enhance early identification of abnormal operations by around 18-25% relative to existing approaches, based on the reported improvements in recall. Also, if they focus their limited resources on processes with a CRS greater than 0.7 (a high-risk level), audit agencies might work better. This might reduce the time it takes to go over risky trades by as much as 30%.

It is now much easier to find fraud in Latin America's public buying processes because LAIRM was formed by the government to do just that. It is designed to be true, clearly measured, and able to respond to different conditions. This turns the research from a deep study into a new, approved research that follows scientific standards.

5. DISCUSSION

The real results show that LAIRM is very helpful for finding data-based risks of fraud among public buyers. These results refer to theory and practice as a whole in this talk. It provides clarity regarding the pros and cons of the study by naming them and specifying how they will affect future policy and research.

5.1. Additions to theory and practice

The LAIRM was created and approved to go after people who steal state contracts, which is a big area of crime. This model only deals with fraud on a small scale, so it is easy to create and check signs. It may be possible to make this work in other parts of the government where theft happens more often.

The study improves the scientific standards of computational social science by formalizing a model that is clear, repeatable, and quantitatively verified. Institutional economics (Acemoglu & Restrepo, 2018) and principal-agent theory (Rajala & Jalonen, 2025) are two ideas that LAIRM draws on. This brings the general ideas about government and real-world policy facts together.

Also, the results show the importance of being able to change with the situation. The model performed better overall, and its F1-score was 15% higher than the World Bank standard (Smidt et al., 2025), but this was not just because of better algorithms. It was also designed specifically for the data infrastructures and corruption modes of Latin America. Changes made after the contract in Mexico, contracting networks in Brazil, and sub-national competition in Colombia, all show that there are important changes in risk factors. This supports the main idea that analytics that work must consider the way institutions are set up in each place. This study questions the idea that uniform models, which are usually made in OECD frameworks, can be easily used in different contexts (Fazekas & Kocsis, 2020).

The LAIRM is not only a study project, but also a helpful tool that can improve the speed and quality of work done by public audit groups. The method may help get past a lack of resources and move from investigations based on scandal to proactive, risk-based tracking with the help of a proven, automatic risk-scoring system.

5.2. What this means for methods and policy

This type of analysis has a big effect on the field of anti-corruption analytics, and the strict measure method sets a new standard for judging new ideas in this field. If future studies develop new tools or signs, they must be compared to current standards in the same way that they are done now, in order to demonstrate that they are better.

The data points to a two-sided method for lawmakers and Supreme Audit Institutions (SAIs) in Latin America. First, modern society needs to invest significantly in data control and sharing, so that the benefit of these models can be maximized. The issues with preparation show that the quality of analytics depends on the quality of the data used. Secondly, the findings support the planned use of risk-based auditing. Using a proven model like LAIRM to find high-risk operations, and then focusing limited investigative resources on those procedures, is a logical and evidence-based way to get the most out of anti-corruption efforts.

5.3. Limitations and avenues for future research

This study, while useful, is limited in a number of ways that show where more research is needed. It is a big problem in today's world to use old tags to verify information. It is difficult to identify when certain actions are honest or not, which could lead to illegal results with criminal consequences.

If SAIs are used currently, they should be able to test the model with real-world data available today. Today, people need to know how to manage technology correctly in order to make more accurate decisions. If the process is biased in any way, it will be easier to verify if it favors certain types of jobs or companies, which will allow us to detect and prevent damages to the stake holders and reduce damage control during the process.

Finally, because industry 5.0 will be focused on people, the next big step is to create a “human-in-the-loop” interface. To achieve this, the best route is to show the risk inspectors how to manage the scores to determine when a process conforms to institutional guidelines.

In summary, the study shows that a context-based analysis model, based on theory and careful testing, can be very helpful in the fight against public buying fraud.

6. CONCLUSIONS

This study approaches a major institutional problem, namely fraud. People developed the LAIRM and set up tests and guidelines to solve these problems and find fraud risk in public buying. The study found that a clear, reliable, and locally suitable mathematical model might make control methods much more accurate and useful compared to the standards used today in every country. LAIRM is based on the principal-agent model and ideas about how businesses work. You can obtain a risk score with the help of signs and weights that give you a head start. Currently, this is a good perspective. Tight cross-validation and sensitivity analysis were both used to test this number.

People today are more likely to believe the plan because there is real-world, verifiable data. The F1 score is 15% higher than the Red Flags method from the World Bank (Smidt et al., 2025) and 9% higher than the model from Fazekas & Kocsis, 2020. Modern society should be able to find what they need. There are various red flags in Mexico, Brazil, and Colombia. The best way to spot risks is to look at how crime works in each country.

In Mexico, for instance, people often switch jobs when their contract expires. In Brazil, on the other hand, people often work in groups. Early spotting rates will go up by 18–25%, according to this model. This shows that it can help the government give people more tools and checks. It talks about some of its flaws, like how it used data that was not supposed to be there, and how hard it was to put together data from different Latin American government websites. It is very costly to handle data and verify its accuracy in the real world because of these problems.

To begin, national audit groups are useful, since they can help confirm that the model is correct and can provide more accurate numbers. Also, there are always new things that need to be changed, such as how fair and open the models are and how bias is measured. This is to ensure the safety of the people watching and those being watched. People who are knowledgeable about politics and business need to work with data scientists to make sure that mixed models are built properly and context is taken into account.

This research shows a practical model that can be modified to fit various scenarios, and is intuitive and user-friendly. People wanting to purchase in Latin America can use it to determine the likelihood of being subject to fraud. Based on the findings, rules, processes, and procedures are proposed to support the model of advanced public contracting management. The LAIRM is the proposed model of this study, being a useful and fact-based tool for stakeholders.

REFERENCES

Acemoglu, D., & Restrepo, P. (2018). Artificial intelligence, automation, and work. In A. Agrawal, J. Gans & A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda* (pp. 197-236). University of Chicago Press. http://www.nber.org/books/agra-1

Adam, I., & Fazekas, M. (2021). Are emerging technologies helping win the fight against corruption? A review of the state of evidence. *Information Economics and Policy*, *57*, 100950. https://doi.org/10.1016/j.infoecopol.2021.100950

Agencia Nacional de Contratación Pública Colombia Compra Eficiente. (2025). *SECOP II*. https://www.colombiacompra.gov.co/secop/secop-ii

Agu, J. C., Nkwo, F. N., & Eneiga, R. U. (2024). Governance and anti-corruption measures in Nigeria: Strategies for enhancing transparency, accountability and public trust. International *Journal of Economics and Public Policy*, *8*(1), 1-15. https://doi.org/10.5281/zenodo.12576796

Alfian, A., Ritchi, H., & Adrianto, Z. (2023). Fraud analytics practices in public-sector transactions: a systematic review. *Journal of Public Budgeting, Accounting & Financial Management*, *35*(5), 685-710. https://doi.org/10.1108/JPBAFM-11-2022-0175

Alnuaimi, N. T., CHatha, K. A., & Abdallah, S. (2024). Role of big data analytics and information processing capabilities in enhancing transparency and accountability in e-procurement applications. *Journal of Engineering, Design and Technology*, *23*(5), 1729-1750. https://doi.org/10.1108/JEDT-12-2023-0544

Androutsopoulou, M., Askounis, D., Carayannis, E. G., & Zotas, N. (2024). Leveraging AI for enhanced eGovernment: Optimizing the use of open governmental data. *Journal of the Knowledge Economy*, *16*, 12998-13033. https://doi.org/10.1007/s13132-024-02317-w

Arellano-Gault, D., Rojas-Salazar, G., & Vargas-Pineda, L. R. (2025). Critical theory in public policy and anti-corruption strategies in Latin America: unveiling the limitations (and hypocrisy?) of the dominant anti-corruption discourse. In L. Secchi & C. N. Cruz-Rubio (Eds.), *Handbook of Public Policy in Latin America* (pp. 356-371). Elgar.

Bahoo, S., Alon, I., & Floreani, J. (2021). Corruption in economics: a bibliometric analysis and research agenda. *Applied Economics Letters*, *28*(7), 565-578. https://doi.org/10.1080/13504851.2020.1764476

Barata, J., & Kayser, I. (2023). Industry 5.0 – Past, Present, and Near Future. *Procedia Computer Science*, *219*, 778-788. https://doi.org/10.1016/j.procs.2023.01.351

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., Borkulo, C. D. van, Bork, R. van, & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 58. https://doi.org/10.1038/s43586-021-00055-w

Breuer, M., Labro, E., Sapra, H., & Zakolyukina, A. A. (2024). Bridging theory and empirical research in accounting. *Journal of Accounting Research*, *62*(3), 1121-1139. https://doi.org/10.1111/1475-679X.12545

Brückner, A., Wölke, M., Hein-Pensel, F., Schero, E., Winkler, H., & Jabs, I. (2025). Assessing industry 5.0 readiness—Prototype of a holistic digital index to evaluate sustainability, resilience and human-centered factors. *International Journal of Information Management Data Insights*, *5*(1), 100329. https://doi.org/10.1016/j.jjimei.2025.100329

Chen, C., & Ganapati, S. (2023). Do transparency mechanisms reduce government corruption? A meta-analysis. *International Review of Administrative Sciences*, *89*(1), 257-272. https://doi.org/10.1177/00208523211033236

Chystiakova, A., Podliehaiev, K., Khoronovskyi, O., Sokur, T., & Kubariev, I. (2025). The role of big data analytics in the investigation of corruption offences. *Revista Jurídica Portucalense*, 116-143. https://doi.org/10.34625/issn.2183-2705(37)2025.ic-6

Controladoria Geral da União. (2025). *Portal da Transparência*. https://portaldatransparencia.gov.br/

Davis, D. B., Mendoza, R. U., & Yap, J. K. (2024). Corruption risk and political dynasties: exploring the links using public procurement data in the Philippines. *Economics of Governance*, *25*(1), 81-109. https://doi.org/10.1007/s10101-023-00306-4

Dessureault, J.-S., Lamontagne, R., & Parisé, P.-O. (2025). The ethics of creating artificial superintelligence: a global risk perspective. *AI and Ethics*, *5*, 6241-6263. https://doi.org/10.1007/s43681-025-00793-7

Didier, N. (2025). Decolonizing Public Administration in Latin America: A systematic literature review of trending discussions in the region. *Public Administration and Development*, 1-18. https://doi.org/10.1002/pad.2092

Fazekas, M., & Kocsis, G. (2020). Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data. *British journal of political science*, *50*(1), 155-164. https://doi.org/10.1017/S0007123417000461

Genest, Y. (2025). The evolution of internal audit in anti-corruption activities: leveraging data analytics and it technology. *EDPACS*, *70*(4), 60-66. https://doi.org/10.1080/07366981.2025.2453271

Ghobakhloo, M., Iranmanesh, M., Mubarak, M. F., Mubarik, M., Rejeb, A., & Nilashi, M. (2022). Identifying industry 5.0 contributions to sustainable development: A strategy roadmap for delivering sustainability values. *Sustainable Production and Consumption*, *33*, 716-737. https://doi.org/10.1016/j.spc.2022.08.003

Gobierno de México. (2025). *Compras MX*. https://comprasmx.buengobierno.gob.mx/

Gutiérrez-Romero, R., & Iturbe, N. (2024). Causes and electoral consequences of political assassinations: The role of organized crime in Mexico. *Political Geography*, *115*, 103206. https://doi.org/10.1016/j.polgeo.2024.103206

Horochovski, R. R., Mancuso, W. P., & Pereira, T. F. (2024). A produção acadêmica sobre o Caso Lava Jato. *Revista de Sociologia e Política*, *32*, e021. https://doi.org/10.1590/1678-98732432e021

Kaswan, M. S., Chaudhary, R., Garza-Reyes, J. A., & Singh, A. (2024). A review of Industry 5.0: from key facets to a conceptual implementation framework. *International Journal of Quality & Reliability Management*, *42*(4), 1196-1223. https://doi.org/10.1108/IJQRM-01-2024-0030

Kenny, C., & Musatova, M. (2010). Red flags of corruption in World Bank projects: An analysis of infrastructure contracts. In S. Rose-Ackerman & T. Søreide (Eds.), *International handbook on the economics of corruption* (Vol. 2, pp. 299-335). Edward Elgar Publishing. https://doi.org/10.1596/1813-9450-5243

Kirby, N. (2021). An 'institution-first'conception of public integrity. *British journal of political science*, *51*(4), 1620-1635. https://doi.org/10.1017/S000712342000006X

Klitgaard, R. (1988). *Controlling corruption*. University of California Press.

Kour, R., Karim, R., Dersin, P., & Venkatesh, N. (2024). Cybersecurity for Industry 5.0: trends and gaps. *Frontiers in Computer Science*, *6*, 1434436. https://doi.org/10.3389/fcomp.2024.1434436

León, A. J. A. V. (2012). *Revisión de las reformas a la Ley de Adquisiciones, Arrendamientos y Servicios del Sector Público del 28 de mayo de 2009, en materia de licitaciones públicas* [Undergraduate thesis, Centro de Investigación y Docencia Económicas]. Repositorio Digital CIDE. http://hdl.handle.net/11651/4829

Linhartova, V., & Pucek, M. J. (2024). Corruption and Human Development: Panel Data Analysis in Transition Economies. *Montenegrin Journal of Economics*, *20*(2), 169-182. https://doi.org/10.14254/1800-5845/2024.20-2.14

Luthra, A., Dixit, S., Garg, S., Singh, A., & Anchal, M. (2025). Addressing Ethical Considerations and Responsible AI Practices. In A. Behl, C. Krishnan, P. Malik & S. Gautam (Eds.), *The ChatGPT Revolution: How Conversational AI is Transforming Customer Service and Business Operations* (pp. 129-149). https://doi.org/10.1108/978-1-83549-852-120251007

Madhala, R. T. (2025). Driving Digital Transformation in Insurance: The Role of System Integration and Guidewire Expertise. *International Journal of Emerging Trends in Computer Science and Information Technology*, 91-100. https://doi.org/10.56472/ICCSAIML25-111

Meyer, K. Z., Luiz, J. M., & Fedderke, J. W. (2025). Corruption dynamics: Integrating structure, agency and institutional logics across contexts. *International Journal of Management Reviews*, *28*(1), e12403. https://doi.org/10.1111/ijmr.12403

Mittal, P. (2020). Big data and analytics: a data management perspective in public administration. *International Journal of Big Data Management*, *1*(2), 152-165. https://doi.org/10.1504/IJBDM.2020.112415

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

North, D. C. (2025). Institutions and the Performance of Economies Over Time. In C. Ménard & M. M. Shirley (Eds.), *Handbook of New Institutional Economics* (pp. 25-35). Springer. https://doi.org/10.1007/978-3-031-50810-3_2

Organisation for Economic Co-operation and Development. (2019). *Government at a Glance 2019*. OECD Publishing. https://www.oecd.org/en/publications/government-at-a-glance-2019_8ccf5c38-en.html

Rabuzin, K., & Modrusan, N. (2019, September). Prediction of Public Procurement Corruption Indices using Machine Learning Methods. *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Vienna, Austria. https://doi.org/10.5220/0008353603330340

Rajala, T., & Jalonen, H. (2025). Beyond Simplification in Public Sector Accountability: Contradictions Between Principal-Agent and Complexity Theories. *Public Administration Review*, *85*(6), 1757-1770. https://doi.org/10.1111/puar.13941

Rojas, V. C. J. (2022). *Responsabilidad precontractual del estado: imprecisiones en el sistema de subsanabilidad en Colombia a partir del SECOP II*. [Undergraduate thesis]. Universidad Externado de Colombia.

Rose-Ackerman, S., & Palifka, B. J. (2016). *Corruption and government: Causes, consequences, and reform*. Cambridge University Press.

Smidt, H., Mitchell, N. J., & Bakke, K. M. (2025). A red flag for public goods? The correlates of civil society restrictions. *Governance*, *38*(2), e12894. https://doi.org/10.1111/gove.12894

Souza, P. V. N. C. S., Ramos, T. M., & Silva, L. G. (2021). Inclinações pragmáticas na nova Lei de Licitações e Contratos Administrativos (Lei nº 14.133/2021): novos princípios, velhos problemas. *Revista de Direito Brasileira*, *29*(11), 4-15. https://doi.org/10.26668/IndexLawJournals/2358-1352/2021.v29i11.7514

Storey, V. C., Baskerville, R. L., & Kaul, M. (2025). Reliability in design science research. *Information Systems Journal*, *35*(3), 984-1014. https://doi.org/10.1111/isj.12564

Trequattrini, R., Palmaccio, M., Turco, M., & Manzari, A. (2024). The contribution of blockchain technologies to anti-corruption practices: A systematic literature review. *Business Strategy and the Environment*, *33*(1), 4-18. https://doi.org/10.1002/bse.3327

Velásquez, A. G., & Díez, C. A. D. (2019). El contrato de prestación de servicios en el estatuto general de contratación de la administración pública: un tipo contractual de compleja celebración. *Vniversitas*, *68*(139), 1-17. https://doi.org/10.11144/Javeriana.vj139.cpse

Wachs, J., & Kertész, J. (2019). A network approach to cartel detection in public auction markets. *Scientific Reports*, *9*(1), 10818. https://doi.org/10.1038/s41598-019-47198-1

Wahab, S., Imran, M., Ahmed, B., Rahim, S., & Hassan, T. (2024). Navigating environmental concerns: Unveiling the role of economic growth, trade, resources and institutional quality on greenhouse gas emissions in OECD countries. *Journal of Cleaner Production*, *434*, 139851. https://doi.org/10.1016/j.jclepro.2023.139851

Yatsyna, Y., & Kudinov, I. (2023). Innovative analytical and statistical technology as a corruption counteraction tool: conceptual analysis. *Amazonia Investiga*, *12*(67), 78-86. https://doi.org/10.34069/AI/2023.67.07.7

Gabriel Silva

Ph.D. in Business Management from Instituto Tecnologico de Costa Rica; Research at Universidad Latinoamericana de Ciencia y Tecnología (ULACIT). E-mail: gsilvaa468@ulacit.ed.cr

AUTHORS' CONTRIBUTION

Gabriel Silva-Atencio: Conceptualization (Lead); Data curation (Lead); Formal analysis (Lead); Investigation (Lead); Methodology (Lead); Project administration (Lead); Resources (Lead); Software (Lead); Validation (Lead); Visualization (Lead); Writing – original draft (Lead); Writing – review & editing (Lead).

RESEARCH DATA AVAILABILITY STATEMENT

The entire dataset supporting the results of this study was published in the article itself.

CONFLICT OF INTEREST

The author has no conflicts of interest to declare.

ARTIFICIAL INTELLIGENCE USAGE

The Grammarly artificial intelligence tool was used to assist with the technical and grammatical review of the text and the standardization of references.

ACKNOWLEDGEMENTS

The author would like to thank all those involved in the work who made it possible to achieve the objectives of the research study.