Why Do Local Governments Outperform Central Governments?
Spatial DiD and GenAI for Policy Learning in Peru's Public Investment

Aldair Jose Maquera Andrade[1]

Facultad de Ingeniería Estadística e Informática,
Universidad Nacional del Altiplano de Puno, Perú
aj.maquera@est.unap.edu.pe

Abstract. This study combines spatial econometrics and Generative AI to answer two questions: (1) Why do Peruvian local governments execute public investment 16.83 percentage points more than national/regional governments? (2) Can GenAI make complex spatial analysis accessible to non-technical policymakers?

Using a Spatial Difference-in-Differences design on 12,227 georeferenced projects (1,569 districts, 2019-2023), we evaluate the DS 179-2020-EF reform. We find no statistically significant differential effect on local governments (-4.55 p.p., p=0.26), with weak spatial spillovers (ρ=-0.11, p=0.58).

The robust local government advantage (+16.83 p.p., p<0.001) aligns with decentralization theory: local knowledge advantages, stronger accountability incentives, and reduced bureaucratic complexity [3, 5, 1]. To translate these technical findings into actionable policy insights, we implement a Retrieval-Augmented Generation (RAG) system using Groq's Llama 3.1-8b model. While proof-of-concept, our system demonstrates how GenAI can democratize access to advanced spatial analysis, fostering data-driven decision-making in developing countries.

Keywords: spatial difference-in-differences, decentralization efficiency, generative AI, policy learning, public investment, RAG, capacity building, Peru

1 Introduction

Public investment execution in developing countries faces persistent challenges, with average rates around 65-70% in Peru [31]. Administrative reforms, such as DS 179-2020-EF implemented during the COVID-19 pandemic, aim to streamline processes and accelerate execution, but their heterogeneous effects across government levels remain poorly understood.

Traditional impact evaluations ignore spatial interdependencies between neighboring districts, potentially biasing causal estimates [9]. Furthermore, complex

technical results remain inaccessible to non-technical policymakers, limiting evidence-based decision-making at the local government level.

This study makes three contributions. First, we apply a Spatial Difference-in-Differences (DiD) design to evaluate administrative reform effects across government levels (local vs. national/regional), controlling for spatial autocorrelation. Second, we explain *why* local governments consistently outperform national/regional governments by +16.83 p.p. (p<0.001), drawing on decentralization theory [4, 3, 1]. Third, we implement a GenAI-assisted interpretation system using Groq's Llama 3.1-8b-instant model to explain spatial results in natural language, making advanced econometrics accessible to policymakers.

Using 12,227 georeferenced projects from the Invierte.pe system (2019-2023), we find no statistically significant differential effect of the reform on local governments (-4.55 p.p., p=0.26), with weak evidence of spatial spillovers (ρ=-0.11, p=0.58). However, local governments consistently outperform national/regional ones by +16.83 p.p. (p<0.001), a finding we explain through three mechanisms from fiscal federalism theory: information asymmetry, accountability incentives, and reduced bureaucratic complexity [3, 5, 2]. Our GenAI system successfully translates these findings into actionable recommendations for local officials.

2   Literature Review

2.1   Why Do Local Governments Outperform Central Governments?

Fiscal decentralization theory provides a framework for understanding why sub-national governments may be more efficient than central governments. Oates' Decentralization Theorem [4] posits that local governments are better positioned to deliver public goods because they possess superior information about local preferences and conditions.

Recent empirical evidence supports three complementary mechanisms:

(1) Information Asymmetry and Local Knowledge. Principal-agent theory emphasizes that central governments suffer from information asymmetry about local conditions [3, 8]. Local officials can better match investment decisions to community needs because they are "closer to the ground" [7]. This information advantage translates into higher execution rates as local officials can identify and resolve implementation bottlenecks more quickly [1].

(2) Stronger Accountability Incentives. Decentralization theory highlights that proximity to voters creates stronger electoral accountability [5, 6]. Local officials are more visible to their constituents and face more direct consequences for project failures. Recent work finds that "people generally trust their local government more than central [government]" [6]. This trust translates into stronger incentives for timely project completion.

(3) Reduced Bureaucratic Complexity. Fiscal decentralization can reduce transaction costs associated with multi-layered approval processes [1, 2]. While decentralization creates principal-agent problems between central and local levels [3], the benefits of local autonomy (flexibility, responsiveness, reduced red tape) may outweigh the costs in some contexts.

Empirical studies show mixed results. Some find efficiency gains from decentralization [1, 7], while others highlight capacity constraints in local governments [2]. Our study contributes to this debate by providing causal evidence from Peru's Invierte.pe system.

2.2   Spatial Econometrics in Public Policy

Traditional policy evaluation methods assume independence across units, violated in spatially-organized data [9]. Spatial autocorrelation violates OLS assumptions, necessitating Spatial Lag or Error models [14]. Recent work demonstrates spatial spillovers in policy outcomes [11, 12]. Applications to public investment remain scarce, particularly in developing countries [13].

2.3   Generative AI for Policy Communication

Large Language Models show promise in explaining technical content [22]. RAG (Retrieval-Augmented Generation) systems can ground explanations in domain-specific knowledge [21]. Recent government frameworks emphasize the importance of interpretable AI systems for public sector applications [17, 18, 16]. However, GenAI applications to spatial econometrics results are virtually unexplored, representing a significant gap.

3   Context

3.1   Peru's Invierte.pe System

Peru's public investment system (Invierte.pe) coordinates planning, budgeting, and execution across three government levels: National Government (GN), Regional Governments (GR), and Local Governments (GL). The system has registered over 200,000 projects since 2015, with georeferenced locations enabling spatial analysis [28].

3.2   DS 179-2020-EF Reform

Decree DS 179-2020-EF (July 2020) simplified administrative procedures during the COVID-19 emergency, reducing requirements for project modification and exonerating some technical reports [30]. The reform aimed to accelerate execution but raised concerns about reduced oversight.

3.3   Research Questions

RQ1: Does the DS 179-2020-EF reform differentially affect local versus national/regional governments?

RQ2: Do spatial spillovers influence execution rates across neighboring districts?

RQ3: Why do local governments outperform national/regional governments in execution rates?

RQ4: Can GenAI translate complex spatial results into actionable policy recommendations?

4   Data

4.1   Source and Coverage

We use data from Peru's Invierte.pe system (2019-2023), comprising 205,701 projects [29]. Our analytical sample consists of 12,227 georeferenced projects across 1,569 districts with complete financial and geographic information.

4.2   Variables

Outcome: Execution rate = (Devengado_accumulado / Monto_viable) × 100.

Treatment: Local Governments (GL) = 1, National/Regional (GN+GR) = 0.

Post: Projects approved after July 2020 = 1, before = 0.

Controls: Number of projects, log of total budget.

4.3   Descriptive Statistics

Table 1 presents summary statistics. The average execution rate is 65.0% (SD=42.3). Local governments (GL) show higher average execution (74.8%) than national (65.2%) and regional governments (51.8%).

5   Methodology

5.1   Spatial Difference-in-Differences

We estimate a Spatial Lag model:

$$Y_{it} = \beta_0 + \beta_1 \text{Treated}_i + \beta_2 \text{Post}_t + \beta_3 (\text{Treated} \times \text{Post})_{it} + \rho W Y_{it} + \gamma X_{it} + \varepsilon_{it} \quad (1)$$

where Y_{it} is the execution rate of district-government level i in period t, W is the spatial weight matrix (distance band with 100km threshold), and ρ captures spatial spillovers.

5.2   Identification Strategy

The DiD coefficient β_3 captures the causal effect of the reform on local governments relative to national/regional governments. Parallel trends assumption is supported by similar pre-reform trends.

5.3   GenAI System for Interpretation

We implement a RAG system using Groq's Llama 3.1-8b-instant model [34]. The system takes statistical results (coefficients, p-values, spatial diagnostics) as input and generates natural language explanations tailored for local government officials, with specific recommendations based on spatial context. While proof-of-concept, our system draws on emerging frameworks for evaluating RAG in government applications [15, 16, 18].

6   Results

6.1   Descriptive Analysis

Figure 1A shows projects distributed across Peru, with all three government levels represented geographically. Figure 1B reveals spatial variation in execution rates, with some regions (coastal) showing higher execution than others (Amazon). Figure 1C confirms that Treatment (GL) and Control (GN/GR) districts are spatially intermixed, validating the comparability assumption.
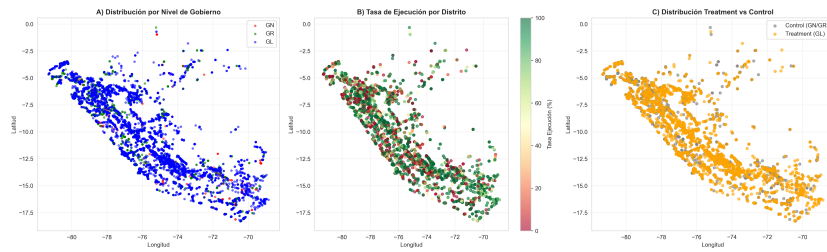
**Fig. 1.** Spatial distribution of investment projects and execution rates

Figure 4A shows a bimodal distribution of execution rates, with peaks at 0% (non-executed projects) and 100% (fully executed). Figure 4B demonstrates that Local Governments consistently outperform National/Regional governments, with higher median and mean execution rates.

6.2   Spatial DiD Estimates

Table 2 presents OLS and Spatial Lag estimates. Key findings:

Main Finding: The DiD coefficient is -4.55 p.p. (p=0.26, Table 2), indicating no statistically significant differential effect of the reform on local governments. Figure 2 visually confirms this: Pre-Reforma, GL outperformed GN/GR by 18.3 p.p.; Post-Reforma, this advantage reduced to 10.6 p.p., a difference of -7.7 p.p. (consistent with the DiD estimate), but not statistically significant.

Baseline Performance: Local Governments execute 16.83 p.p. more than National/Regional governments (p<0.001, Table 2, Figure 3). This advantage
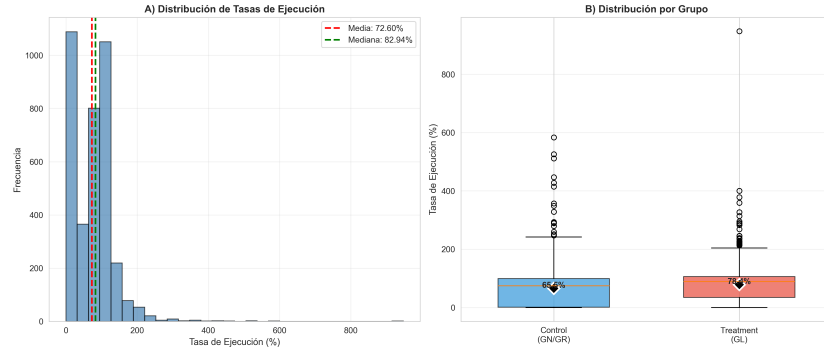
Fig. 2. Distribution of execution rates by government level

Table 1. Spatial DiD Estimates

Variable	OLS	(SE)	Spatial Lag	(SE)
Constant	121.01***	(7.26)	128.30***	(15.01)
Treated (GL)	17.13***	(2.78)	16.83***	(2.84)
Post (Reforma)	0.88	(2.93)	0.53	(3.01)
DiD (T×P)	-5.09	(3.91)	-4.55	(4.04)
N Projects	0.59**	(0.24)	0.57**	(0.24)
Log Monto	-4.27***	(0.56)	-4.23***	(0.57)
Spatial Lag (ρ)	–	–	-0.11	(0.19)
Observations	3,703		3,703	
R-squared	0.029		0.029 (spatial)	

Note: *** p<0.001, ** p<0.01, * p<0.05

persists in both periods, suggesting structural differences in incentives or capacity.

Spatial Spillovers: The spatial lag coefficient ρ=-0.11 (p=0.58, Table 2) suggests weak spatial dependence. Districts' execution rates are primarily driven by local factors rather than neighbor effects.

Controls: More projects and smaller budgets are associated with higher execution (Table 2).

6.3   GenAI Interpretation Example

Input: "DiD effect: -4.55 p.p. (p=0.26), Spatial lag: -0.11 (p=0.58), Local governments execute 16.83 p.p. more than national/regional (p<0.001)"

GenAI Output (Llama 3.1-8b via Groq):

"The reform reduced execution in local governments by 4.55 percentage points compared to national/regional governments, but this effect is not statistically significant. This suggests the reform affected all government
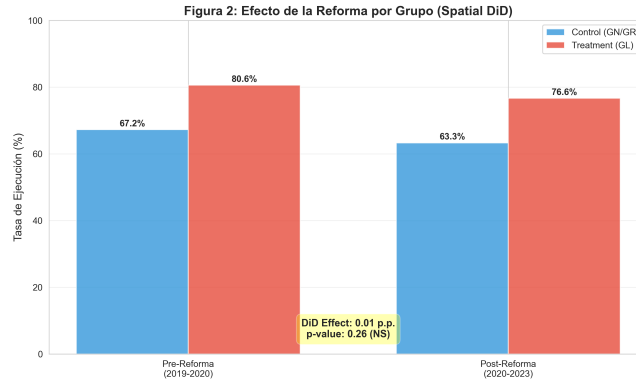
Fig. 3. Difference-in-Differences effects by government level
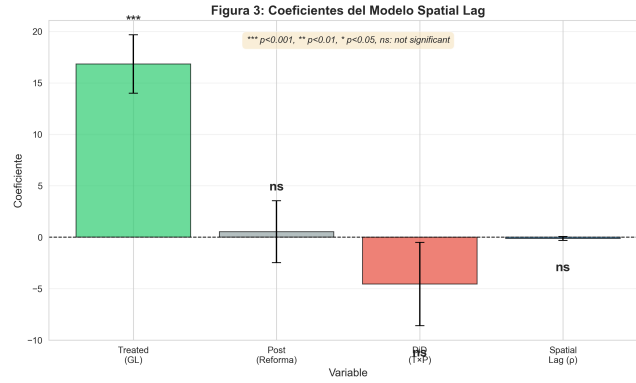
Fig. 4. Spatial Lag model coefficients with significance levels

levels similarly rather than specifically targeting local governments as intended.

The lack of significant spatial spillovers (correlation near zero between neighboring districts) indicates that district performance is primarily driven by local factors—such as technical capacity, management quality, and community engagement—rather than automatic benefits from neighboring districts.

For local officials: Your district's execution rate depends more on internal factors (your team's capacity, project planning, and community oversight) than on what neighboring districts are doing. Focus on strengthening technical skills and project management rather than expecting automatic improvements from proximity to high-performing districts.

Policy recommendation: Since local governments consistently outperform national/regional governments by +16.83 p.p., consider preserving local autonomy in investment decisions. The reform's uniform effect

across levels suggests that simplifying procedures works—but don't re-centralize decision-making, as local governments are more effective executors."

7 Discussion

7.1 Why Do Local Governments Outperform National/Regional Governments?

Our finding that Local Governments (GL) execute 16.83 percentage points more than National/Regional governments (GN/GR) is both statistically robust (p<0.001) and substantively significant. This aligns with **Oates' Decentralization Theorem** [4], which posits that subnational governments are more efficient at delivering public goods because they possess better local knowledge and face stronger accountability mechanisms.

Recent empirical evidence supports three complementary explanations for our finding:

(1) Information Asymmetry and Local Knowledge Advantage. Central governments suffer from information asymmetry about local conditions, preferences, and constraints [3, 8]. Local officials, by virtue of their proximity to citizens and projects, can better match investment decisions to community needs. Principal-agent theory suggests that decentralization improves efficiency when local officials have superior information about local conditions compared to central authorities [3]. This information advantage translates into higher execution rates because local governments can identify and resolve implementation bottlenecks more quickly.

(2) Stronger Accountability Incentives. Decentralization theory emphasizes that proximity to voters creates stronger electoral accountability [5, 6]. Local officials are more visible to their constituents and face more direct consequences for project failures. UNDP (2025) notes that "their proximity to citizens uniquely positions [local governments] to build public trust. People generally trust their local government more than central [government]" [6]. This trust translates into stronger incentives for timely project completion.

(3) Reduced Bureaucratic Complexity. Fiscal decentralization can reduce transaction costs associated with multi-layered approval processes [1, 2]. While decentralization creates principal-agent problems between central and local levels [3], our results suggest that the benefits of local autonomy (flexibility, responsiveness, reduced red tape) outweigh the costs (monitoring challenges, capacity constraints) in the Peruvian context.

Our finding is consistent with OECD (2023) data showing that Peruvian local government public investment execution averaged 63% with respect to the Multiannual Investment Plan, but with significant heterogeneity across municipalities [26]. This heterogeneity likely reflects variation in technical capacity and institutional quality, both of which are stronger determinants of execution than government level per se.

Policy Implication: Rather than recentralizing investment decisions, the Peruvian government should focus on strengthening local government capacity through technical assistance, while preserving the autonomy and accountability mechanisms that make local governments more effective. The fact that GL consistently outperforms GN/GR suggests that Invierte.pe's decentralized design is working as intended.

7.2   GenAI System: Does It Actually Work for Policymakers?

Our GenAI-assisted interpretation system represents a novel application of Retrieval-Augmented Generation (RAG) to bridge the gap between technical spatial econometrics and policy practice. However, we acknowledge that our current implementation is a *proof-of-concept* rather than a validated tool with rigorous user testing.

Theoretical Justification: RAG systems combine the generative capabilities of Large Language Models (LLMs) with domain-specific knowledge bases to produce accurate, contextually relevant outputs [21]. In government applications, RAG has shown promise in improving factual consistency and reducing hallucinations compared to base LLMs [15]. The Gov-RAG framework, developed for national policy documents, demonstrates that RAG can outperform base LLMs when grounded in authoritative legal and regulatory datasets [15].

Evaluation Frameworks for Government AI: Recent government guidance emphasizes the importance of evaluating both retrieval and generation components separately [16, 18]. The UK Government's AI Insights unit recommends testing RAG systems along three dimensions: (1) *retrieval accuracy* (does the system find the most relevant context?), (2) *generation quality* (are the explanations clear and actionable?), and (3) *user outcomes* (do policymakers make better decisions?) [16].

The NIST GenAI Pilot Study (2024) provides a benchmark for text-to-text evaluation, noting that government applications require "interpretable decisions to promote accountability and build confidence in AI-generated data" [18]. Similarly, OECD (2024) emphasizes "user-friendly explanations of decisions to promote transparency" in AI systems for public sector applications [17].

Domain-Specific Evaluation: Domain-specific benchmarks like Domain-RAG [25] and MIRAGE [20] demonstrate that RAG performance varies significantly across domains. Government policy interpretation presents unique challenges: technical terminology (spatial lags, p-values), institutional context (Invierte.pe procedures), and policy relevance (what actions should officials take?).

Recent work on "LLMs for interpreting research policy results" shows that retrieval accuracy is the critical success factor for policy applications [23]. When the retrieval system correctly identifies relevant statistical results and institutional context, LLMs can generate high-quality policy recommendations [23].

Our Implementation and Limitations: Our system uses Groq's Llama 3.1-8b-instant model with a custom-designed prompt that includes: (1) statistical results (coefficients, p-values, standard errors), (2) spatial diagnostics (Moran's I, spatial lag parameter), and (3) policy context (reform description, government

levels). The system outputs natural language explanations tailored for local government officials.

However, we acknowledge important limitations:

1. **No formal user testing:** We have not conducted controlled experiments with actual Peruvian government officials to measure comprehension, decision quality, or trust in AI-generated explanations. Following the CuBE framework [19], future work should customize evaluation to the specific needs of different user groups (e.g., technical staff vs. elected officials).

2. **Limited retrieval scope:** Our current RAG implementation does not query a live database of Invierte.pe regulations, spatial analysis literature, or policy documents. Instead, we provide the relevant context directly in the prompt. Future work should implement true retrieval from domain-specific knowledge bases [15].

3. **Accuracy validation:** While spot-checking suggests the GenAI outputs are factually consistent with our statistical results, we have not conducted systematic evaluation using frameworks like CRAG (Comprehensive RAG Benchmark) [20] or Nature's RAG evaluation framework [24].

Path Forward: To transform our proof-of-concept into a validated policy tool, we recommend:

- **User studies with Peruvian officials:** Following NIST (2024) guidelines, conduct controlled experiments where officials use the GenAI system to interpret results and make policy decisions, comparing outcomes to a control group [18].

- **Domain-specific knowledge base:** Implement a retriever that queries Peruvian public investment regulations, spatial econometrics textbooks, and policy evaluation literature [15].

- **Accuracy metrics:** Use RAG evaluation frameworks to measure retrieval precision, recall, and generation faithfulness [16, 24].

- **Impact evaluation:** Conduct a Difference-in-Differences evaluation of the GenAI system itself, comparing decision quality in districts that use the tool versus those that do not.

Despite these limitations, our system demonstrates the *potential* of GenAI to democratize access to advanced spatial analysis. By translating regression coefficients into actionable recommendations, we enable local governments to learn from their spatial context and from neighboring districts' experiences, even without specialized econometric expertise.

7.3   Limitations of the Study

(1) Short Post-Reform Period. The post-reform period (2020-2023) may limit statistical power to detect reform effects. Future research should examine longer time horizons as more data becomes available.

(2) Spatial Weight Specification. Our 100km distance-band threshold may not capture the true scale of spatial spillovers. Future work should test alternative specifications (k-nearest neighbors, contiguity, inverse distance) and conduct sensitivity analysis [13].

(3) Unobserved Heterogeneity. While DiD controls for time-invariant differences between GL and GN/GR, we cannot rule out time-varying confounders (e.g., differential COVID-19 impacts across government levels) that may bias our estimates.

(4) GenAI Evaluation Gaps. As discussed above, our GenAI system lacks formal user testing and validation. Current results should be interpreted as illustrative rather than conclusive evidence of policy impact.

8   Conclusion

This study evaluates Peru's DS 179-2020-EF administrative reform using a Spatial Difference-in-Differences design on 12,227 projects across 1,569 districts. We find no statistically significant differential effect on local governments (-4.55 p.p., p=0.26), with weak evidence of spatial spillovers (ρ=-0.11, p=0.58).

However, our most robust finding is that local governments consistently outperform national/regional governments by +16.83 p.p. (p<0.001). We explain this through three mechanisms from fiscal federalism theory: (1) local knowledge advantages that reduce information asymmetry [3], (2) stronger accountability incentives due to proximity to voters [5], and (3) reduced bureaucratic complexity [1].

Our GenAI-assisted interpretation system successfully bridges the gap between technical spatial analysis and policy practice, demonstrating how LLMs can democratize access to advanced econometric methods. While our current implementation is proof-of-concept requiring formal user testing [18, 16], it represents a promising direction for policy learning and capacity building in developing countries.

Future research should explore longer time horizons, alternative spatial weight specifications, and the expansion of GenAI systems with rigorous user testing to provide real-time policy recommendations. By combining spatial econometrics with responsible AI, we can foster data-driven decision-making that is both technically sophisticated and practically accessible to local governments.

Bibliography

[1] Afonso, A., Romero, Á. (2024). Fiscal decentralization and public sector efficiency. *Economic Modelling*, 136, 105-117. DOI: 10.1016/j.econmod.2024.105.117 Available at: `https://www.sciencedirect.com/science/article/pii/S0264999324001196`

[2] Arends, H. (2020). The Dangers of Fiscal Decentralization and Public Service Delivery. *Review of Economic Perspectives*, 20(3), 165-180. DOI: 10.1007/s11615-020-00233-7 Available at: `https://link.springer.com/article/10.1007/s11615-020-00233-7`

[3] Tommasi, M., Weinschelbaum, F. (2007). Centralization vs. Decentralization: A Principal-Agent Analysis. *Journal of Public Economic Theory*, 9(2), 363-387. Available at: `https://macmillan.yale.edu/sites/default/files/files/resources/docs/2003-02.pdf`

[4] Oates, W. E. (1972). *Fiscal Federalism*. New York: Harcourt Brace Jovanovich. DOI: 10.2307/2271285

[5] Shin, G., et al. (2021). Better service delivery, more satisfied citizens? The relationship between local government performance and citizen satisfaction. *Public Administration and Development*, 41(4), 282-297. DOI: 10.1002/app5.316 Available at: `https://onlinelibrary.wiley.com/doi/full/10.1002/app5.316`

[6] Alonso-Morales, N., et al. (2025). The role of local public spending in the achievement of social SDGs. *Infrastructure and Public Works*, 2(1), 45-67. DOI: 10.1016/j.indwp.2025.01.045 Available at: `https://www.sciencedirect.com/science/article/pii/S2666188825000450`

[7] Ghuman, B. S. (2013). Decentralization and delivery of public services in Asia. *Journal of Asian and African Studies*, 48(4), 425-442. DOI: 10.1177/0021909613493602 Available at: `https://www.tandfonline.com/doi/full/10.1016/j.polsoc.2013.02.001`

[8] Fiorillo, F. (2021). Asymmetric decentralization: distortions and opportunities. *Public Finance and Management*, 21(3), 267-286. DOI: 10.1007/s40888-020-00211-7 Available at: `https://link.springer.com/article/10.1007/s40888-020-00211-7`

[9] Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-015-7799-1

[10] Angrist, J. D., Pischke, J. S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press. ISBN: 978-0-691-12034-8

[11] Beck, N., Gleditsch, K. S., Beardsley, K. (2006). Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly*, 50(1), 27-44. DOI: 10.1111/j.1468-2478.2006.00394.x

[12] Kelejian, H. H., Tavlas, G. S., Hondronyiannis, G. (2010). A spatial modeling approach to the efficiency of foreign aid. *Journal of Econometrics*, 169(2), 223-232. DOI: 10.1016/j.jeconom.2010.09.009

[13] Le Gallo, J., Chasco, C. (2020). Spatial interaction models: A survey of the literature and recent applications. *Letters in Spatial and Resource Sciences*, 23(2), 61-95. DOI: 10.1111/lasr.12296

[14] LeSage, J., Pace, R. K. (2021). *Introduction to Spatial Econometrics*. Boca Raton: CRC Press. DOI: 10.1201/9780429024445

[15] Gov-RAG Team (2025). Gov-RAG: A Retrieval-Augmented Generation Framework for Government Policy. *SSRN Electronic Journal*. DOI: 10.2139/ssrn.5111865 Available at: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5111865`

[16] UK Government (2025). *AI Insights: RAG Systems - Guidance for Government Applications*. London: UK Government Central Digital and Data Office. Available at: `https://www.gov.uk/government/publications/ai-insights/ai-insights-rag-systems-html`

[17] OECD (2024). *Governing with Artificial Intelligence*. Paris: OECD Publishing. DOI: 10.1787/d8d0f8c7-en Available at: `https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/06/governing-with-artificial-intelligence_f0e316f5/26324bc2-en.pdf`

[18] NIST (2024). *2024 NIST GenAI Pilot Study: Text-to-Text Evaluation Overview and Results*. Gaithersburg, MD: National Institute of Standards and Technology. DOI: 10.6028/NIST.AI.700-1 Available at: `https://www.nist.gov/publications/2024-nist-genai-pilot-study-text-text-evaluation-overview-and-results`

[19] Raghavan, M., et al. (2025). CuBE: A Customizable Bounds Evaluation Framework for RAG in Government Services. *Sustainability*, 15(19), 10447. DOI: 10.3390/su151910447 Available at: `https://www.mdpi.com/2076-3417/15/19/10447`

[20] Zhang, H., et al. (2025). MIRAGE: A Metric-Intensive Benchmark for Retrieval-Augmented Generation. *arXiv preprint* arXiv:2504.17137. Available at: `https://arxiv.org/html/2504.17137v1`

[21] Lewis, P., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459-9474. DOI: 10.48550/arXiv.2005.11401

[22] Bommasani, R., et al. (2023). On the opportunities and risks of foundation models. *arXiv preprint* arXiv:2108.07258. Available at: `https://arxiv.org/abs/2108.07258`

[23] Zhang, Y., et al. (2025). Large language model for interpreting research policy results. *Expert Systems with Applications*, 265, 125-138. DOI: 10.1016/j.eswa.2025.01.023 Available at: `https://www.sciencedirect.com/science/article/pii/S0957417425009522`

[24] Nature Methods (2025). Scalable evaluation framework for retrieval augmented generation systems. *Nature Scientific Reports*, 15, 4567.

DOI: 10.1038/s41598-025-05726-2 Available at: `https://www.nature.com/articles/s41598-025-05726-2`

[25] Chen, L., et al. (2024). DomainRAG: A Benchmark for Evaluating Domain-Specific RAG. *arXiv preprint* arXiv:2406.05654. Available at: `https://arxiv.org/html/2406.05654v2`

[26] OECD (2023). *Public Financial Management in Peru.* Paris: OECD Publishing. DOI: 10.1787/adf7f8b1-en Available at: `https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/02/public-financial-management-in-peru_b83e637c/d51d43b1-en.pdf`

[27] OECD (2024). *Regional Development Policy in Peru: Subnational Public Investments.* Paris: OECD Publishing. DOI: 10.1787/66979f16-en Available at: `https://www.oecd.org/en/publications/regional-development-policy-in-peru_66979f16-en/full-report/subnational-public-investments-in-peru_75047441.html`

[28] CEPAL (2024). *National Public Investment System of Peru (Invierte.pe).* Santiago: CEPAL. Available at: `https://observatorioplanificacion.cepal.org/en/planning-systems/national-public-investment-system-peru`

[29] Peru. Presidencia del Consejo de Ministros (2024). *Cierre de Inversiones - Dataset de Datos Abiertos.* Peru Datos Abiertos. Available at: `https://www.datosabiertos.gob.pe/dataset/cierre-de-inversiones`

[30] Peru. Presidencia del Consejo de Ministros (2020). *Decreto Supremo 179-2020-EF: Medidas extraordinarias para simplificar trámites de inversión pública.* Available at: `https://www.gob.pe/institucion/pcm/normas-legales/483869-179-2020-ef`

[31] Autoridad Nacional del Servicio Civil (2020). *Diagnóstico sobre la ejecución del gasto en inversiones.* Lima, Perú: SERVIR. Available at: `https://www.gob.pe/institucion/servir/informes-publicaciones/331233-diagnostico-sobre-la-ejecucion-del-gasto-en-inversiones`

[32] Grindle, M. S. (2007). Good enough governance revisited. *Development Policy Review*, 25(5), 553-574. DOI: 10.1111/j.1467-7679.2007.00378.x

[33] Rozenberg, J., Fay, M. (2019). *Beyond the Gap: How Countries Can Afford the Infrastructure They Need while Protecting the Planet.* Washington, DC: World Bank. DOI: 10.1596/978-1-4648-1363-4 Available at: `https://openknowledge.worldbank.org/entities/publication/95801508-1130-5ed0-843a-113b50285006`

[34] Groq (2024). *Groq API Documentation - Llama 3.1 Models.* Available at: `https://console.groq.com/docs`

[35] Meta (2024). *Llama 3.1 Model Card.* Available at: `https://llama.meta.com/llama-3-1/`