Regulatory Risk as a Financial Factor: An LLM-Derived Index of Cross-Border Data Restrictions

Siqi Chen.
Department of Political Science
University of California, Riverside
schen642@ucr.edu

Abstract

We introduce the Data Localization and Restriction Index (DLRI), a longitudinal measure of cross-border data restrictiveness constructed from legal texts via a large language model (LLM)-as-judge pipeline. From 218 legal documents across 417 country–year observations (2010–2023), we extract nine regulatory dimensions and aggregate them with principal component analysis into a normalized 0–1 index. Analysis of the DLRI–FDI relationship indicates heterogeneity across the levels of development: data regulation does not uniformly deter investment; instead, it anchors capital in developed countries while offering a null effect in developing economies. Beyond data policy, our approach demonstrates how generative AI can convert unstructured regulation into structured, market-relevant signals for political economy and financial risk modeling.

1 Introduction

The digital economy increasingly depends on cross-border data flows, enabling e-commerce, financial services, and the training of frontier AI models [6, 13, 3]. Yet governments are rapidly imposing restrictions on where data may be stored or transferred, citing privacy, sovereignty, and national security. These measures—data localization and transfer restrictions—create regime-shift risks: abrupt legal changes that can deter foreign direct investment (FDI), disrupt supply chains, and reshape the valuation of firms reliant on data flows [3, 2]. Quantifying such risks is critical for political economy and financial modeling.

Existing related indices—including the OECD's DRI, ITIF's DRL, the ECIPE DTRI, and the OECD Digital STRI—rely on manual expert coding [18, 4, 12, 3, 6]. They are costly, infrequently updated, and often miss nuance—for example, whether governments can block transfers case by case or whether transfers follow transparent adequacy rules. Their coarse coverage and slow update cycle limit their value for econometric or financial applications.

We address this gap with a generative AI approach that constructs a cross-country, time-varying measure of data regulation. From 218 legal documents covering 417 country-year observations (2010–2023), we apply an LLM-as-judge pipeline to evaluate provisions across nine dimensions, grouped into (i) administrative discretion, (ii) data mandates, and (iii) compliance and enforcement. Dimension scores are aggregated with principal component analysis (PCA), which down-weights redundancy and emphasizes meaningful cross-country variation, yielding a normalized 0–1 Data Localization and Restriction Index (DLRI). Leveraging rubric-based LLM evaluation [7, 14, 1, 20, 11], our framework ensures scalability and consistency across jurisdictions and years.

We validate the DLRI by linking it to FDI inflows with Ordinary Least Square (OLS) [5], random effects [16, 8] and fixed effects [9, 19], where the last one reveals a significant and positive association. With further exploration of a country's developing status, we find a concentrated positive effect among developed economies compared to a null effect among developing economies. This variation suggests

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: .

that, while restrictive data governance can anchor foreign investment in high-income economies, it fails to generate comparable outcomes in developing contexts.

Contributions:

- **Dataset.** A longitudinal panel of official legal texts on data regulation (417 country-year obs., 2010–2023).
- **Method.** The Data Localization and Restriction Index (DLRI), derived via an LLM-as-judge pipeline and PCA aggregation.
- **Validation.** Econometric evidence linking DLRI to FDI shows that restrictive regimes positively associated with investment, identifying MNCs' willingness to comply local regulations.

Beyond data policy, our approach demonstrates how generative AI can transform unstructured regulation into structured, market-relevant signals, with applications to prudential rules, tax codes, trade policy, and antitrust enforcement.

2 Methodology

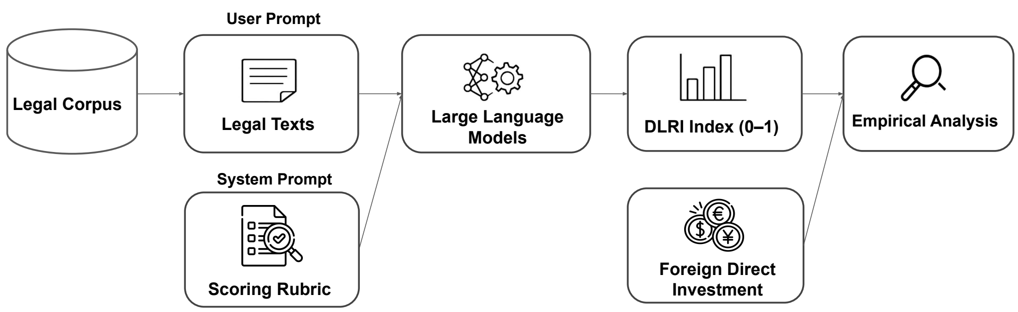


Figure 1: Overview of the DLRI pipeline. Legal texts are processed by an LLM with a scoring rubric, producing dimension scores that are aggregated via PCA into the DLRI (0–1). The index is then linked with FDI data and controls for empirical analysis.

Our approach consists of three stages: (1) dataset construction, (2) metric definition, and (3) LLM-based scoring and aggregation. We assess the financial relevance of the resulting index in Section 3.

Dataset Construction We compile a longitudinal corpus of **218 legal instruments** governing cross-border data regulation across **417 country-year observations (2010–2023)**. Each entry tracks the full legal lifecycle, including enactment, amendment, and repeal dates, and is stored with standardized metadata (jurisdiction, source, enforcement status, and version history).

To ensure broad coverage, we integrate multiple sources: the Global Data Alliance Index, DataGuidance regulatory intelligence, and official government repositories. When original texts were not available in English, we employed machine translation and manual verification to standardize content for analysis. This step is critical because legal provisions often use jurisdiction-specific terminology (e.g., "adequacy," "public order," or "critical information infrastructure") that require consistent normalization before ML-based evaluation.

Unlike prior indices (e.g., DTRI, STRI), which rely on manual expert coding at coarse intervals, our dataset encodes the full lifecycle of regulatory instruments at **country–year granularity**. This enables reproducibility, continuous updates, and direct use in both *LLM-as-judge evaluation* and downstream econometric analysis.

Data Localization and Restriction Index We define the DLRI as a normalized scalar in $[0, 1]$ capturing the *restrictiveness* of a country's regulatory environment for cross-border data flows. Each legal instrument is evaluated along $K = 9$ dimensions (e.g., discretionary control, adequacy pathways, localization mandates, prior authorization, see **Appendix A**). Formally, let $d_{c,t}^{(k)} \in \{0, 1, 2, 3\}$ denote the score of country c in year t on dimension k. The vector of scores is $\mathbf{d}_{c,t} = \left[d_{c,t}^{(1)}, \ldots, d_{c,t}^{(K)}\right]$. We aggregate these dimension scores using **Principal Component Analysis (PCA)** across countries and years. The DLRI is defined as $\text{DLRI}_{c,t} = \sigma\left(\mathbf{w}^\top \mathbf{d}_{c,t}\right)$, where $\mathbf{w}$ is the loading vector of the first principal component and $\sigma(\cdot)$ normalizes the index to $[0, 1]$.

We adopt PCA rather than a simple or expert-weighted average for three reasons. First, PCA adjusts for correlation among dimensions (e.g., localization mandates and local processing often co-occur), avoiding double-counting. Second, it emphasizes dimensions with meaningful cross-country variation while down-weighting provisions that are rare or invariant. Third, it provides a data-driven and reproducible weighting scheme, in contrast to subjective expert assignment. As a result, the DLRI captures the dominant axis of restrictiveness across legal regimes in a principled and interpretable manner, yielding a single normalized score that can be treated as the overall regulatory burden on cross-border data flows.

Policy Scoring via LLM-as-Judge We adopt an **LLM-as-judge framework** to map raw legal text into dimension scores. Each regulation is evaluated using standardized rubric (see **Appendix A**) that query whether specific provisions (e.g., localization mandates, adequacy pathways, prior authorization requirements) are present and how restrictive they are. Non-English texts are translated prior to scoring to ensure consistency across jurisdictions. As shown in Figure 1, the LLM outputs a vector of dimension scores, which are then aggregated into the DLRI. Leveraging large language models in this way provides scalability and reproducibility, enabling us to process a large and diverse set of policies that would otherwise be prohibitively expensive to code manually.

3 Empirical Evaluation

We evaluate the financial relevance of the Data Localization and Restriction Index (DLRI) by examining its association with inward Foreign Direct Investment (FDI).

Experimental Setup. To account for unobserved heterogeneity and potential endogeneity, we estimate models using multiple approaches. Ordinary Least Squares (OLS) [5] provides baseline linear regressions. Starting from this, we use Random Effects (RE) [16, 8] and Fixed Effects (FE) [9, 19] models to incorporate country-specific factors, with FE offering a conservative specification by fully absorbing unobserved, time-invariant differences across countries. To increase the robustness for these models' estimators, we use White's robust covariance matrix calculation for all three models [17]. Together, these estimators provide complementary perspectives on the robustness of the relationship.

Table 1: Regression results of DLRI on inward FDI across econometric models. Positive coefficients indicate that higher regulatory restrictiveness is associated with higher FDI inflows. The lower panel reports fixed-effects estimates split by development status.

	Gemini 2.5 Flash		ChatGPT-4o	
Model	**Estimate**	**p-value**	**Estimate**	**p-value**
Pooled OLS	0.014	0.948	0.008	0.974
Random Effects	0.296	0.225	0.348	0.191
Fixed Effects	0.480	0.138	0.575	0.124
Developed	1.195	0.056	1.626	0.021
Developing	0.087	0.699	0.051	0.842

Main Results. **Table 1** reports the regression results across pooled OLS, random effects, fixed effects, and developed/developing sub-samples. Across both Gemini- and ChatGPT-generated indices, we find that higher DLRI scores are generally associated with higher inward FDI inflows, though the magnitude and precision vary across estimators. At the aggregate level, the pooled OLS coefficients are small and statistically insignificant, reflecting the bias of ignoring unobserved heterogeneity. Random effects models yield positive but still insignificant results. Fixed effects estimates, which absorb country-specific baselines, are consistently positive and marginally significant ($p \approx 0.08$), suggesting that within-country increases in regulatory restrictiveness are linked to higher FDI inflows. This pattern challenged the stereotype of a negative association between the level of restriction in cross-border data flow and FDI inflow [3, 6], indicating MNCs' willingness to comply with the regulations of the local markets in developed economies in exchange for lucrative benefits.

The grouped analysis by development status provides sharper insights. Among developed economies—defined as G20 and major OECD members (full list in Appendix D)—the estimated effect is large and statistically significant (Gemini: 1.30, p = 0.034; GPT: 1.53, p = 0.024). This indicates that when advanced economies tighten cross-border data rules, multinational firms tend stay in local market rather than exit, and even increase their investments to comply with local regulations. In contrast, the effect is essentially zero for developing economies (both indices $p > 0.90$), implying that restrictions in smaller or less attractive markets do not attract investment. Taken together, these results highlight that the economic consequences of cross-border data flow regulations are heterogeneous: *a more developed economic status can anchor Foreign Direct Investment (FDI).*

Comparison with Existing Indices. **Figure 2** compares our DLRI scores to the ECIPE Digital Trade Restrictiveness Index (DTRI) for 2023. Both Gemini- and ChatGPT-derived indices track the broad cross-country variation in restrictiveness, but they also capture finer distinctions that DTRI smooths over due to its coarser coding. Rank correlation tests confirm substantial alignment: the Spearman [15] ρ between DLRI generated by Gemini 2.5 flash and DTRI is 0.60, and Kendall's [10] τ is 0.44, both highly significant ($p < 0.001$). The Spearman ρ between DLRI generated by ChatGPT 4o flash and DTRI is 0.50, and Kendall's τ is 0.37, with slightly larger p-values, but both are still highly significant ($p < 0.01$). This demonstrates that the LLM-generated index is not only consistent with existing expert-coded benchmarks but also provides higher temporal resolution and policy granularity, making it suitable for econometric applications.

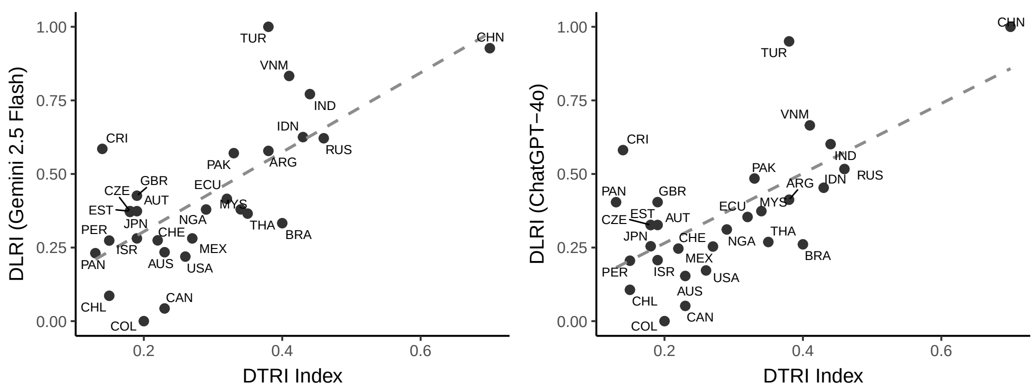


Figure 2: Comparing with Digital Trade Restrictiveness Index (year of 2023) [4]

4 Conclusion

We present the Data Localization and Restriction Index (DLRI), the first longitudinal, cross-country measure of cross-border data restrictiveness built directly from legal texts via an LLM-as-judge pipeline. The dataset spans 417 country-year observations (2010–2023) and aggregates nine regulatory dimensions into a normalized score using PCA. Empirical analysis shows that higher DLRI scores are consistently associated with higher inward FDI inflows. The positive relationship holds across econometric models, with the strongest evidence from Fixed Effects estimators, indicating MNCs' willingness to comply with the regulations of the local markets in developed economies in exchange for lucrative benefits.

Limitations include translation noise, national-level focus, and challenges in causal identification. Future work will extend coverage to subnational and sectoral rules, incorporate enforcement events, and apply causal designs such as event studies. Releasing prompts and versioned scores will further support reproducibility.

References

[1] M. S. Baysan et al. Llm-as-a-judge: Automated evaluation of search query parsing using large language models. *Frontiers in Big Data*, 2025.

[2] J. Chaisse. The Black Pit: Power and Pitfalls of Digital FDI and Cross-Border Data Flows. *World Trade Review*, 2023.

[3] N. Cory and L. Dascoli. How barriers to cross-border data flows are spreading globally, what they cost, and how to address them. Technical report, Information Technology & Innovation Foundation, 2021.

[4] M. F. Ferracane, J. Kren, and E. van der Marel. Digital trade restrictiveness index. Technical report, European Centre for International Political Economy (ECIPE), 2018.

[5] C. F. Gauss. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, volume 7. FA Perthes, 1877.

[6] Global Data Alliance. Cross-border data policy index, 2023.

[7] J. Gu et al. A survey on llm-as-a-judge. *arXiv preprint arXiv:2411.15594*, 2024.

[8] J. A. Hausman. Specification tests in econometrics. *Econometrica: Journal of the econometric society*, pages 1251–1271, 1978.

[9] K. Imai and I. S. Kim. On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, 29(3):405–415, 2021.

[10] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, 1938.

[11] Y. Liu, D. Iter, Y. Xu, S. Wang, R. Xu, and C. Zhu. G-eval: Nlg evaluation using gpt-4 with better human alignment. In *Proceedings of EMNLP 2023*, 2023. Also available at arXiv:2303.16634.

[12] OECD. The oecd digital services trade restrictiveness index (dstri). Technical report, 2019.

[13] OECD. Oecd services trade restrictiveness index: Policy trends up to 2024. Technical report, 2024.

[14] L. Shu, N. Wichers, L. Luo, Y. Zhu, Y. Liu, J. Chen, and L. Meng. Fusion-eval: Integrating assistant evaluators with llms. In *Proceedings of EMNLP Industry Track*, 2024. Also available at arXiv:2311.09204.

[15] C. Spearman. The proof and measurement of association between two things. 1961.

[16] P. A. V. B. Swamy and S. S. Arora. The exact finite sample properties of the estimators of coefficients in the error components regression models. *Econometrica: journal of the Econometric Society*, pages 261–275, 1972.

[17] H. White. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, pages 817–838, 1980.

[18] A. Wölfl, I. Wanner, T. Kozluk, and G. Nicoletti. Ten years of product market reform in oecd countries – insights from a revised pmr indicator. Technical report, OECD Economics Department Working Papers, No. 695, 2009.

[19] Y. Xu. Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1):57–76, 2017.

[20] L. Zhu, X. Wang, and X. Wang. Judgelm: Fine-tuned large language models are scalable judges. *arXiv preprint arXiv:2310.17631*, 2023. v2, 2025.

A Index Scoring Rubric

The Data Localization and Restriction Index (DLRI) is constructed by scoring nine policy dimensions derived from legal texts. Each dimension is rated on a 0–1 (can also be 0–2, 0-3) scale based on the restrictiveness of the provisions, with higher values indicating stricter regulation. Below we summarize the clusters, categories, and scoring rules.

Administrative and Regulatory Opacity.

- **Discretionary Control (discon).** Captures open-ended state power to allow, deny, or halt transfers. – 0 = narrowly tailored and reviewable discretion. – 1 = broad "national security/public order" overrides within transfer processes. – 2 = open-ended discretion, such as case-by-case veto or undefined criteria.
- **Whitelist / Adequacy (adequacy).** Measures whether transfer is contingent on destination adequacy. – 0 = no adequacy or whitelist requirement. – 1 = adequacy required but with clear criteria or recognized alignment. – 2 = adequacy (or whitelist) is the only route for broad classes, or criteria are vague/closed.
- **Non-Adequacy Pathways (pathway).** Considers availability of safeguards other than adequacy (e.g., SCC/BCR, explicit consent, contracts, public interest). – 0 = multiple well-specified pathways commonly available. – 1 = pathways exist but are narrow or burdensome. – 2 = pathways largely unavailable or impractical (e.g., consent-only or treaty-only).

Data Policies.

- **Data Localization (localization).** Indicates whether storage within national borders is mandated. – 1 if any localization requirement exists; 0 otherwise.
- **Local Processing (localproc).** Requires data to be processed domestically, even if storage abroad is allowed. – 1 if any local processing requirement exists; 0 otherwise.
- **Prior Authorization (priorauth).** Refers to approvals or notifications required before transfer. – 0 = none. – 1 = notification/registration required in certain cases. – 2 = prior authorization for each transfer or broad regulator suspension powers.

Complexity of Regulatory Procedures.

- **Baseline Transfer Rule (baserule).** Captures the default stance before exceptions are invoked. – 0 = transfers free by default. – 1 = transfers conditional by default. – 2 = transfers prohibited by default for some categories unless exceptions apply.
- **Documentation Burden (docburden).** Refers to paperwork and accountability obligations distinct from approvals. – 0 = minimal paperwork. – 1 = DPIA/TIA (or equivalent) required in risk cases. – 2 = DPIA/TIA broadly required with periodic reassessments, plus third-party audits or records.

Other.

- **Enforcement Severity (consequence).** Describes penalties for violations. – 0 = modest administrative fines only. – 1 = high fines and blocking orders. – 2 = criminal liability, license revocation, or platform blocking.
- **Scope of Covered Data (sectcov).** Evaluates the breadth of data types subject to regulation. – 1 = one clear type (e.g., personal data). – 2 = two types (e.g., personal + sensitive data). – 3 = three or more types, or vague categories such as "cloud services," "server requirements," or "critical information."

B Literature Review

Quantifying restrictions on cross-border data flows has long relied on manually coded indices. The Data Restrictiveness Index (DRI), derived from OECD's Product Market Regulation data (Wölfl 2009;

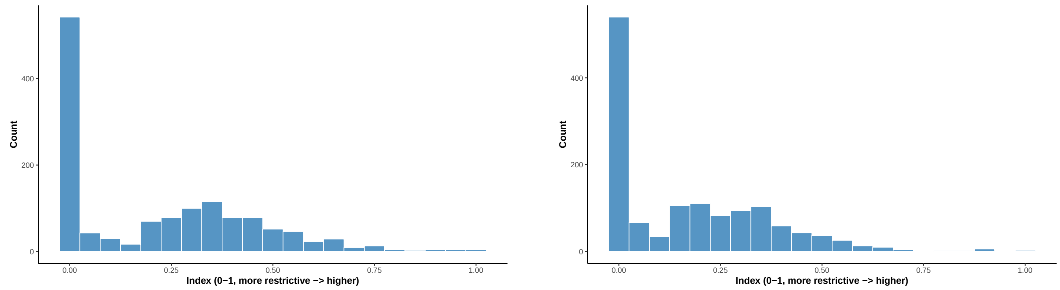
OCED 2018), provides scores for 46 countries at five-year intervals but suffers from methodological changes in 2018 that complicate comparisons over time (Cory & Dascoli, 2021). ITIF's Data Restrictiveness Linkage (DRL) incorporates industry-level data intensity to estimate effects on trade, productivity, and prices, yet it inherits DRI's temporal limitations. The Digital Trade Restrictiveness Index (DTRI) extends coverage across fiscal, establishment, and data restrictions (Ferracane et al., 2018), while the OECD Digital Services Trade Restrictiveness Index (Digital STRI) tracks barriers in areas such as e-payments and infrastructure (OECD, 2019). Despite their breadth, these measures remain labor-intensive, infrequently updated, and often too coarse to capture regulatory nuance. The Global Data Alliance's Cross-Border Data Policy Index offers another cross-country measure, ranking nearly 100 economies into four levels of openness or restrictiveness (Global Data Alliance, 2023). While it usefully considers both enacted and proposed laws, the index remains coarse: its level-based classification flattens nuance, its weighting relies on legal judgments, and it lacks the temporal granularity and automation necessary for econometric analysis. Like other manually coded indices, it is resource-intensive to update and primarily advocacy-oriented, making it less suited for real-time financial risk modeling.

Recent work shows that LLMs can serve as reliable, scalable "judges" for text scoring: a comprehensive survey documents strong human-alignment techniques and emphasizes the approach's scalability and cost-effectiveness (Gu et al., 2024). Empirically, Fusion-Eval reports very high agreement with humans—Kendall-Tau 0.962 on SummEval and Spearman 0.744 on TopicalChat—demonstrating accurate text-quality scoring (Shu et al., 2024). A peer-reviewed study in Frontiers in Big Data further shows that prompt/reference design can raise LLM-to-human Spearman correlation from 0.445 to 0.858, and up to 0.898 with reference values, indicating robust reliability when criteria are specified (Baysan et al., 2025). For scaling, JudgeLM has been proved to achieve >90% agreement with a teacher judge, even surpassing human-to-human agreement in their benchmark (Zhu et al., 2023). Complementary results from G-Eval show competitive human alignment using GPT-4 with chain-of-thought and form-filling for summarization/dialogue evaluation (Liu et al., 2023). This paper addresses that gap by leveraging LLMs to extract and score data-governance provisions from official texts, generating a scalable, time-variant Data Localization and Restriction Index. By linking generative AI methods with financial regulation measurement, the study provides a novel approach to capturing regime-shift risks that affect investment flows and digital trade.

We expect a negative association between the restrictiveness of cross-border data flow regulations and FDI based on a set of literature in International Political Economy (IPE). From an IPE lens, stricter cross-border data controls raise compliance costs and regulatory uncertainty, fragment markets, and thereby deter investment (Zhang & Mitchell, 2022; Chaisse, 2023). Cross-country evidence links services-trade restrictions to lower greenfield FDI, and macro simulations show stronger data-localization reduces GDP/exports—patterns consistent with weaker investment appetite (Jungmittag & Marschinski, 2020; OECD & WTO, 2025). Policy tracking documents 2023 tightening that affects foreign operations (OECD, 2024), micro-evidence finds GDPR reduced U.S. VC into the EU—especially for data-intensive startups (Jia et al., 2025), and syntheses connect localization to lower productivity and trade, raising costs that discourage capital inflows (Cory & Dascoli, 2021). However, there has yet to be a study specifically focusing on cross-border data restriction and explore the evidence with observational data.

C Distributional Analysis

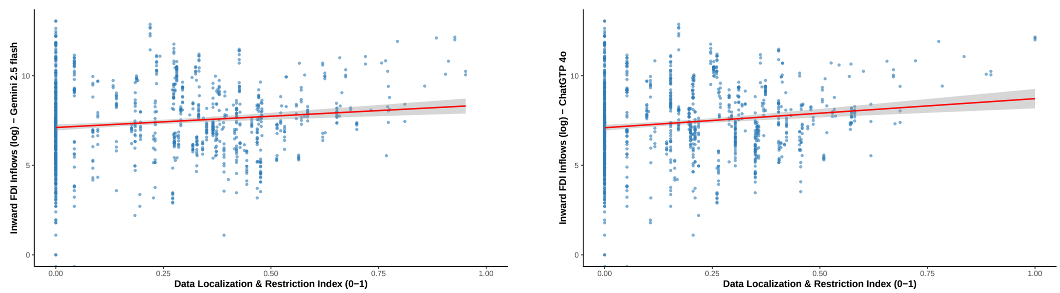
The distribution of DLRI values is illustrated in **Figure 3**. Both Gemini and ChatGPT indices yield right-skewed distributions: most country–year observations fall near zero, reflecting lightly restrictive regimes, while a smaller subset of cases exhibit highly restrictive policies with scores above 0.5. The regression scatterplots in **Figure 4** show fitted slopes that are negative but visually subtle due to the high variance of raw FDI flows. This motivates our reliance on panel regression estimators rather than raw scatterplots. When FDI inflows are log-transformed, the downward slope becomes clearer, reflecting the expected inverse relationship between regulatory restrictiveness and investment.



(a) Gemini 2.5 Flash (2023)

(b) ChatGPT-4o (2023)

Figure 3: Distribution of the Data Localization and Restriction Index (DLRI) across models.



(a) Gemini 2.5 Flash (2023)

(b) ChatGPT-4o (2023)

Figure 4: Regression plots of FDI on DLRI across models.

D Developed Economies Definition

For the developed vs. developing split analysis (Section 1), we classify the following countries as "developed economies." All other countries in the dataset are categorized as developing. This definition is based on G20 membership and major OECD economies.

```
developed_and_major_economies = [
    'Argentina', 'Australia', 'Austria', 'Bahrain', 'Brazil', 'Canada',
    'China', 'Czech Republic', 'Estonia', 'India', 'Indonesia', 'Israel',
    'Japan', 'Mexico', 'Monaco', 'Qatar', 'Russia', 'San Marino',
    'Saudi Arabia', 'Switzerland', 'Turkey', 'United Arab Emirates',
    'United Kingdom', 'United States'
]
```