MAY 2026 | POLICY PROPOSAL

# Artificial Intelligence as a Labor Market Shock: Cross-Country Evidence on National Unemployment

Arjun Paul, Bhoomika Agrawal, Manu Wilson, Mihir Khanna†

# Artificial Intelligence as a Labor Market Shock: Cross-Country Evidence on National Unemployment*

Arjun Paul, Bhoomika Agrawal, Manu Wilson, Mihir Khanna†

May 2026

## Abstract

This paper investigates the effect of AI exposure on national-aggregate unemployment rates across 133 countries from 2010 to 2024. We constructed country-level exposure indices by aggregating occupational exposure scores from Felten et al. (2021) and Eloundou et al. (2023) to the 2-digit ISCO-08 level, using employment shares from ILO data. The correlation between these two indices is very high ($r = 0.99$), indicating that the occupational composition of an economy, rather than the scoring methodology, largely drives country-level exposure. We estimated a 2SLS regression with country- and year fixed effects, instrumenting the interaction between a post-2022 dummy and the country-level FRS index with the Oxford Government AI Readiness Index interacted with the same dummy. The main finding is that more-exposed countries are associated with lower unemployment after 2018: −1.82 percentage points per unit increase in FRS exposure ($p < 0.01$). This finding is robust to alternative occupational scoring, an alternative instrumental variable (the IMF AI Preparedness Index 2023), and the exclusion of the COVID-19 years. In all six specifications, the coefficient on the post-2022 LLM period is positive but not statistically significant. We treat this null finding as indicative rather than conclusive, given that LLM emergence occurred just three years ago, and it agrees with the results of a company-level analysis (Anthropic 2026) that detects no impact until 2025. Overall, the findings suggest that, so far, AI may have affected unemployment mainly through task reallocation and augmentation rather than displacement. It is important to note that, according to the appendix, a pre-trend placebo test suggests that more-exposed countries had a distinctive unemployment trajectory before 2018. As a result, that fact can partly drive the post-2018 coefficients. This paper demonstrates that national unemployment rates cannot detect short-run effects of LLMs and emphasizes the need for future research using firm-, sector-, and task-level data.

**Keywords:** artificial intelligence, labor markets, unemployment, technological change, AI preparedness, instrumental variable, two-stage least squares

* The authors are grateful to Prof. Prachi Mishra and the Isaac Centre for Public Policy (ICPP) team for their valuable suggestions and feedback.

† Isaac Centre for Public Policy (ICPP), Ashoka University, India.

# 1. INTRODUCTION

Artificial intelligence is transforming manufacturing and services by automating routine tasks and augmenting cognitive work. Its adoption occurred in two main stages: the proliferation of pre-large language model (pre-LLM) tools since 2018, and the expansion of LLMs into white-collar and cognitive tasks since 2022. Cross-country adoption of AI has varied, shaped by structural differences in occupational composition, institutional readiness, and absorptive capacity. While there are growing fears regarding technological unemployment, systematic cross-country evidence is sparse, with the vast majority of existing studies being subnational or focused on a single country.

Our analytical framework draws on the literature on labor market shocks. In their seminal paper, Autor et al. (2013) demonstrate that U.S. commuting zones experiencing greater import penetration from China have experienced higher unemployment growth, suggesting that regional structure determines how a shock affects an economy. The shift-share specification, which combines sectoral import penetration ratios with local employment shares, has come to define how labor reallocation due to trade and technology shocks is studied. Another branch of the literature, starting with Rajan and Zingales (1998), considers the role of country absorptive capacity in assessing how the country reacts to an exogenous shock through the interaction of structural variables and country moderators.

This paper extends this line of research to the case of AI adoption, examining the relationship between AI exposure and unemployment across 133 countries from 2010 to 2024. We focus on two distinct stages of AI diffusion: the post-2018 spread of pre-LLMs (transformers, BERT, GPT-1, broader NLP, and computer vision), and the post-2022 spread of LLMs (ChatGPT, Claude, and Gemini). Our research question is, **"how does exposure to AI affect unemployment across countries, and do effects differ between the post-2018 and post-2022 phases?"**

Following Cazzaniga et al. (2024), we combine occupation-level AI scores using the SOC-to-ISCO-08 crosswalk and country-level 2-digit ILO ISCO-08 employment shares. We calculate two such exposure scores using Felten et al. (2021) and Eloundou et al. (2023) scores; these correlate at 0.99 across countries. As the identification method, we use a two-stage least squares regression with fixed effects for country and year, instrumenting AI exposure × Post-2022 with Oxford AI Readiness Index × Post-2022.

We find three main results. First, we observe a negative, statistically significant coefficient on the interaction between AI exposure and the post-2018 variable, which is very robust to instrument choices, scoring methods, and the exclusion of COVID years. Such a result is consistent with the idea that the period after 2018 can be considered one of complementarity, in which AI complements existing tasks and increases workers' productivity. Second, the coefficient on AI exposure interacted with the LLM post-2022 variable is positive but insignificant — likely because we only observe three years of LLM diffusion, for which there is insufficient statistical power; similarly, Anthropic (2026) finds no labor market effect of LLMs as of 2025 using a firm-level approach. Third, pre-trend placebo regressions (see Appendix) indicate that the negative relationship between AI exposure and unemployment exists prior to the sample window, implying structural differences at the country level.

This paper makes three important contributions to the literature. First, it establishes a 133-country measure of AI exposure, using two different occupation-level scores constructed by independent teams. Second, it finds a robust negative relationship between AI exposure and unemployment since 2018. Third, it demonstrates the inadequacy of aggregate unemployment data in measuring effects of LLM diffusion, highlighting the importance of researching AI diffusion at the firm-, sector-, and task-levels.

The rest of the paper is organized as follows: Literature Review (Section 2); Data Sources (Section 3); Methodology (Section 4); Results and Findings (Section 5); Policy Implications (Section 6); Conclusion (Section 7).

## 2. LITERATURE REVIEW

Our identification strategy builds on the shift-share framework first employed by Autor et al. (2013), who define regional exposure to China's import penetration as the inner product of national import penetration and pre-shock regional employment shares. They find that regions with higher import penetration exhibit higher unemployment rates. Likewise, Acemoglu and Restrepo (2020) define exposure to robotics as a similar inner product, with national robotic penetration replacing import penetration, again resulting in significant job and wage losses. In both cases, the strategy is applied to regions within nations and includes an important limitation often ignored: any pre-existing structure may point different regions in different directions before the shock. In concrete terms, our design replaces import penetration with the AI exposure of the occupation, and pre-shock regional shares with ISCO-08 employment shares. We make two important contributions to this design. First, we extend it from a sub-national to a cross-country scope. This helps us account for aggregate-level outcomes that remain out of reach for sub-national analysis. Second, we deal with the aforementioned pre-trend problem by conducting placebo tests (Appendix A2).

Furthermore, the literature on occupation-based AI exposure has grown. Felten et al. (2021) map AI skills on O*NET work abilities and produce the continuous score — the AIOE — for SOC-10 occupations. Eloundou et al. (2023) developed an equivalent index based on LLM skills. We present results using both, demonstrating that the construction of a continuous score does not determine the effect since AIOE was constructed well before LLMs.

The closest precedent to our methodology comes from Cazzaniga et al. (2024), who construct a continuous measure of occupational exposure and then aggregate it at the national level using ISCO-08 employment shares across over 140 countries. For firms, Acemoglu et al. (2022) and Anthropic (2026) use AI-related labor market data collected in vacancies and conversation logs; such information emerges well before any effect is visible in aggregate unemployment rates. Following the method introduced by Cazzaniga et al., we independently validate the exposure measure using two occupation-level continuous scores and explicitly test for pre-trend issues.

## 3. DATA SOURCES

Our dataset combines four sources of information into a cross-sectional time series spanning 2010–2024 for 133 countries. Our dependent variable is the harmonized unemployment rate (% of the labor force), which is derived from the ILO model estimate in the WDI series SL.UEM.TOTL.ZS. The macroeconomic control variables are GDP per capita, population, service/industry share of GDP, inflation rate (annual), and real GDP growth. Both population and GDP per capita variables are used in their logged forms.

Our explanatory variables consist of two country-level AI exposure indices, obtained from occupation-level scores provided by Felten, Raj, and Seamans (2021) and Eloundou et al. (2023), respectively. Occupation scores are converted into ISCO-08 two-digit sub-major groups using the SOC-to-ISCO classification crosswalk from Hardy, Keister, and Lewandowski (2018). Occupation scores are weighted by employment shares for each country in 2010, based on the ILO definition.

$$\text{Exposure}_c = \sum_{o \in O} s_{c,o}^{2010} \cdot \text{score}_o$$

The index is built in three steps. First, each occupation receives an AI exposure score: the FRS AIOE score, or the Eloundou task score E1 + 0.5 · E2. Second, SOC occupation scores are mapped to ISCO-08 2-digit groups using the Hardy, Keister, and Lewandowski (2018) crosswalk, taking a simple mean when several SOC codes fall into a single ISCO group. Third, for each country, we weight these group scores by the country's 2010 ILO employment shares and sum across groups; here, $s_{c,o}$ is country c's 2010 employment share in group o, and $\text{score}_o$ is that group's exposure score. The two indices correlate at 0.99 across 133 countries using both the Pearson and Spearman measures, indicating robustness to the scoring system. Appendix A0 gives the full details.

We use the IMF AI Preparedness Index 2023 (Cazzaniga et al. 2024) as an alternative instrument; its construction and the corresponding robustness results are in Appendix A3.4. Summary statistics for all variables, observation counts by series, and the construction of the estimation sample are reported in Appendix Table A0.

## 4. METHODOLOGY

We use two-stage least squares estimation (2SLS) instead of ordinary least squares estimation (OLS) because an OLS regression of unemployment on AI exposure would give us biased coefficient estimates for several reasons. Firstly, our measure of AI exposure is chosen by firms, and its adoption by any firm is determined by many factors, such as sufficient capital, a skilled labor force, and firms' incentives to automate; simultaneously, these variables have an impact on unemployment. Thus, the exposure interaction is not exogenous, so an OLS regression would incorrectly interpret the correlation between the selection mechanism and the dependent variable as causality. Secondly, there can be omitted variable bias caused by the influence of institutional development, regulation, and demand shocks in particular sectors on both occupational structure (our exposure measure) and unemployment. Since our regressions account for country fixed effects but not their time-varying components, we still need an instrument. Finally, there is a problem with measurement error in our AI exposure index, which results from constructing the index using occupation-level scores from the SOC-to-ISCO-08 crosswalk (Hardy, Keister, and Lewandowski 2018) and then weighting these occupations by their employment shares before the shock occurs.

The Oxford Government AI Readiness Index solves all these problems. The index is used to measure institutional, technical, and infrastructural readiness to adopt the technology — variables that affect the diffusion process but should influence unemployment through AI alone once macroeconomic variables and fixed effects are included in the regression. In all specifications, the first-stage Kleibergen-Paap F-statistic is much larger than 100 (ranging from 106 to 191), well above the Stock-Yogo (2005) critical value of 16.38. We additionally check the exclusion restriction by replacing the Oxford AI Readiness Index with the IMF AI Preparedness Index (Cazzaniga et al. 2024), a different measure of readiness. This exercise produces coefficients of the same signs and similar significance levels, and therefore we recover the effect of AI exposure on unemployment using 2SLS.

We estimate the 2SLS regression where AI exposure interacted with the post-2022 LLM indicator is treated as an endogenous variable, and the Oxford Government AI Readiness Index interacted with the same indicator is used as an instrument. To simplify the notation, we denote the AI exposure trend prior to LLMs introduction as follows:

$$\text{AITrend}_{ct} \equiv \text{Exposure}_c \times \text{Post2018}_t \qquad Z_{ct} \equiv \text{Exposure}_c \times \text{LLM}_t$$

Our first-stage equation is:

$$Z_{ct} = \pi\,(\text{Oxford}_c \times \text{LLM}_t) + \delta\,\text{AITrend}_{ct} + \gamma' X_{ct} + \mu_c + \lambda_t + \nu_{ct}$$

The endogenous regressor $Z_{ct}$ is instrumented by $\text{Oxford}_c \times \text{LLM}_t$, where $\text{Oxford}_c$ denotes the country's Government AI Readiness Index score for 2024. The $\text{AITrend}_{ct}$ term enters as an exogenous control representing the pre-LLM AI diffusion trend. Our second-stage equation is:

$$U_{ct} = \beta_1\,\text{AITrend}_{ct} + \beta_2\,\widehat{Z}_{ct} + \gamma' X_{ct} + \mu_c + \lambda_t + \varepsilon_{ct}$$

In these equations, $U_{ct}$ is the ILO unemployment rate in country c at year t; $\text{Exposure}_c$ is the country-level AI exposure index, computed once per country from pre-shock employment shares; $\text{Post2018}_t$ and $\text{LLM}_t$ equal one from 2018 and 2022 onward, respectively; $X_{ct}$ is a vector of time-varying macroeconomic controls (log GDP per capita, log population, services share of GDP, industry share of GDP, inflation, and GDP growth); $\mu_c$ and $\lambda_t$ are country and year fixed effects; and $\hat{Z}_{ct}$ is the first-stage fitted value of the endogenous regressor. Standard errors cluster at the country level. The coefficient of interest is $\beta_2$, which identifies the causal effect of LLM-era AI exposure on unemployment, while $\beta_1$ captures the average post-2018 diffusion effect. Country fixed effects absorb the time-invariant level of $\text{Exposure}_c$; identification of $\beta_1$ and $\beta_2$ comes from the within-country variation generated by the interactions with $\text{Post2018}_t$ and $\text{LLM}_t$.

Our identification assumptions are set out in Appendix A1: the country and year fixed effects, the instrument's exclusion restriction, and the rationale for instrumenting the post-2022 interaction.

# 5. RESULTS AND FINDINGS

## 5.1 Main Results

Table 1 presents the 2SLS results for six specifications that differ in their exposure index, instrument choice, and sample. The coefficient on the interaction between AI Exposure and the post-2018 dummy is negative and significant across all six columns: countries with high AI exposure have lower unemployment during the post-2018 diffusion period.

The post-2018 coefficient in the main specification, using the FRS index and Oxford instrument with the 133-country panel, is estimated at −1.819 ($p = 0.002$) (column A3). If we use the Eloundou index (column B3), the result is −10.647 ($p = 0.002$). This discrepancy is purely scale-dependent, as the FRS index is normalized to unit variance, whereas the Eloundou index lies on a [0, 1] scale expressed in probability units. Once standardized, the two coefficients are comparable, implying that AI Exposure reduces the unemployment rate by approximately 0.6 percentage points per standard deviation increase in AI Exposure. In other words, the coefficient indicates that pre-LLM AI technologies (NLP, transformers, BERT, GPT-1, and computer vision) affect the labor market by increasing task productivity, not by completely taking over jobs. This interpretation is supported by Anthropic (2026), which finds no impact of LLMs on firms' labor productivity through 2025.

In contrast, the post-2022 results are drastically different. The coefficient on the exposure × LLM-2022 interaction is statistically indistinguishable from zero across all six specifications. Point estimates range from 0.704 to 1.067 in the FRS panel and from 4.173 to 6.148 in the Eloundou panel, with t-statistics between 0.96 and 1.27; the headline specification (A3) yields 0.869 with a standard error of 0.895, implying a 95% confidence interval of [−0.90, 2.64]. None of the combinations of the exposure index, instrument, and sample delivers a significant coefficient. These findings are explained by a lack of statistical power: only three years had elapsed since the emergence of LLMs, limiting the time horizon of cross-country annual data. The results of the first-stage diagnostics indicate that the instrument remains strong across all six specifications (see Appendix A3.1).

The null hypothesis seems to hold regardless of specification. All six specifications represent a single base-case design, with independent variation in three design choices: two exposure indices, two instruments, and two samples. The coefficient on the post-2022 interaction remains insignificant in all six cases. Thus, there is no specification in which the post-2022 estimate would have a p-value less than 0.2. The post-2018 coefficient, in contrast, has t-statistics exceeding 2.7 in absolute value in all specifications, implying sufficient power to uncover a mid-run effect. Had there been a true negative effect of half the magnitude of the post-2018 coefficient, at least one specification would have detected it. Instead, the estimate changes its sign, from −1.82 (post-2018, FRS+Oxford) to +0.87 (post-2022, the same specification). Hence, one can conclude that LLMs have not yet affected the labor market (the former interpretation) or that their impact does not manifest in cross-country aggregate unemployment (the latter interpretation). Both explanations receive further support from Anthropic (2026), which finds no evidence of LLM-driven employment effects among U.S. companies as of 2025, and from Cazzaniga et al. (2024), whose exposure framework likewise implies that LLM effects surface first in disaggregated, occupation-level data rather than in aggregate unemployment.

**Table 1: AI Exposure and Unemployment: IV (2SLS) Estimates**

| | (A1) | (A2) | (A3) | (B1) | (B2) | (B3) |
|---|---|---|---|---|---|---|
| | FRS × IMF AIPI 122 c. | FRS × Oxford 122 c. | FRS × Oxford 133 c. | Elou. × IMF AIPI 122 c. | Elou. × Oxford 122 c. | Elou. × Oxford 133 c. |
| AITrend ($\beta_1$) | −1.608*** | −1.744*** | −1.819*** | −9.516*** | −10.251*** | −10.647*** |
| | (0.580) | (0.582) | (0.590) | (3.200) | (3.227) | (3.306) |

| | | | | | | |
|---|---|---|---|---|---|---|
| $\hat{Z}_{ct}$ ($\beta_2$) | 0.704 | 1.067 | 0.869 | 4.173 | 6.148 | 5.059 |
| | (0.737) | (0.895) | (0.895) | (3.986) | (4.834) | (4.856) |
| ln GDP p.c. | −1.331* | −1.329* | −1.463** | −1.372** | −1.367** | −1.503** |
| | (0.685) | (0.686) | (0.648) | (0.689) | (0.690) | (0.651) |
| ln Population | 15.143*** | 15.253*** | 13.912*** | 14.945*** | 15.059*** | 13.708*** |
| | (3.308) | (3.339) | (3.228) | (3.320) | (3.350) | (3.243) |
| Services share | 0.064 | 0.064 | 0.055 | 0.064 | 0.063 | 0.054 |
| | (0.043) | (0.043) | (0.040) | (0.043) | (0.043) | (0.040) |
| Industry share | 0.033 | 0.032 | 0.031 | 0.034 | 0.033 | 0.032 |
| | (0.046) | (0.046) | (0.043) | (0.045) | (0.046) | (0.043) |
| GDP growth | −0.053* | −0.053* | −0.052** | −0.053* | −0.053* | −0.052** |
| | (0.028) | (0.028) | (0.024) | (0.028) | (0.028) | (0.024) |
| Observations | 1,786 | 1,786 | 1,937 | 1,786 | 1,786 | 1,937 |
| KP F-stat | 174.0 | 105.9 | 110.7 | 190.5 | 117.7 | 119.5 |
| Country/Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes: IV (2SLS) estimates with country and year fixed effects and macroeconomic controls. The endogenous regressor (AI exposure × Post-2022) is instrumented by readiness × Post-2022, using the IMF AI Preparedness Index in the AIPI columns and the Oxford Government AI Readiness Index in the Oxford columns. Columns A1, A2, B1, and B2 use the 122-country sample for which the IMF index is available; columns A3 and B3 use the full 133-country sample identified by the Oxford instrument, whose broader coverage adds the 11 countries the IMF index omits. Moving A1→A2 (and B1→B2) holds the sample fixed and swaps the instrument; moving A2→A3 (and B2→B3) holds the Oxford instrument fixed and expands the sample. Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

## 5.2 Robustness: Alternative Exposure Measures and Instruments

The two exposure indices correlate at 0.99 across the 133 countries despite using different occupational scoring systems, so the magnitude gap between Panels A and B reflects index scale, not measurement disagreement. Columns A1 and A2 vary the instrument holding the index fixed; A1 and B1 vary the index holding the instrument fixed. Figure 1 plots all six estimates.

## 5.3 Robustness: Dropping COVID-19 Years (2020–21)

Removing the COVID-19 period (2020–2021) increases the post-2018 coefficient to −2.567 (FRS) and −14.791 (Eloundou), while the post-2022 coefficient remains statistically insignificant. This suggests that the year fixed effects absorbed pandemic-driven cross-country differences.

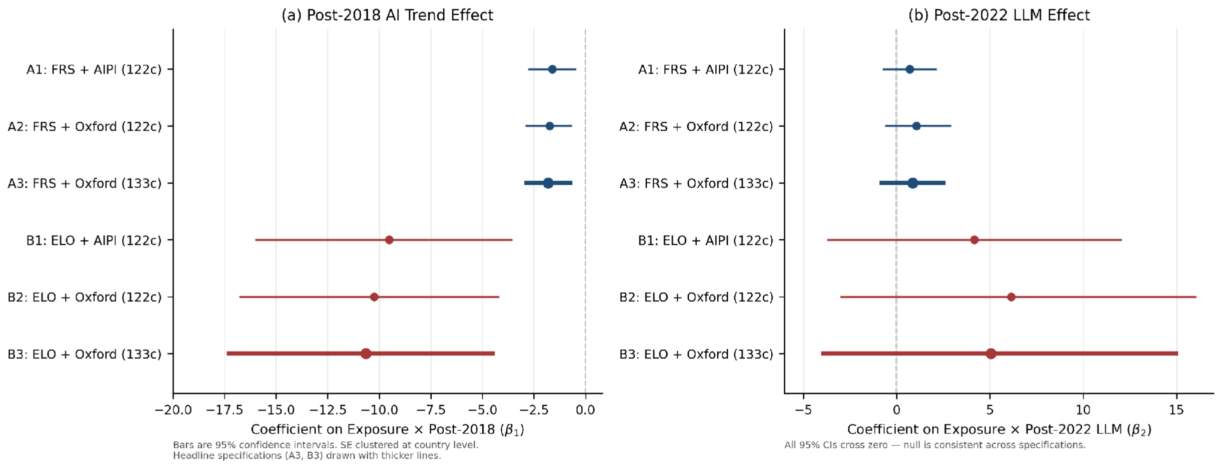


**Figure 1: AI Exposure Effects on Unemployment Across Specifications**

### 5.4 Pre-Trend Caveat

A typical worry about the shift-share approach to exposure measurement is that the cross-sectional component could be correlated with pre-treatment country trends. We test this issue directly by regressing changes in unemployment from 2014–16 to 2017–19 on each of our variables; all placebo regressions are statistically significant at the $p < 0.001$ level. The result shows that countries with high AI exposure/Readiness had lower unemployment rates even before 2018. In other words, the coefficient in question captures not only the impact of AI but also some structural differences, making it an upper bound for the true effect. The relationship in question is informative for policymakers but not prescriptive: it shows what should be invested into adjustment capacity without attributing any unemployment declines to AI alone. Full results of placebo regressions are in Appendix A2.

## 6. POLICY IMPLICATIONS

Four implications can be drawn from the results for government actions concerning the diffusion of AI into the labor market.

First, there is little empirical support for pre-emptive labor-displacement policies. Countries with greater AI exposure experienced decreases in their aggregate unemployment rate during 2018–2024, as shown by the significantly negative post-2018 coefficient. This means that, at the macroeconomic level, the diffusion of AI has led to labor reallocation rather than displacement. However, this does not mean that the transition was costless. Recent research at the firm level shows that, even if aggregate employment rates remain unchanged, AI adoption affects labor hiring and skills (Acemoglu et al., 2022; Anthropic, 2026). Thus, policies should minimize adjustment costs in particular occupations and sectors, for instance, by providing wage insurance, making benefit portability easier, and adopting active labor-market programs to reskill.

Second, optimal AI strategies may need to differ across countries, because exposure and readiness are only weakly aligned (Appendix Figure A4). Singapore and Luxembourg are high on both; Burundi and Ethiopia are low on both. India, by contrast, combines low exposure with high readiness, and this off-diagonal spread makes uniform policy prescriptions inappropriate. High-exposure countries should focus on rapid intra-occupational labor reallocation of skilled workers, including mid-career reskilling, easing cross-sector mobility via reforming occupational licenses, and promoting modular credentialing as a form of partial qualification. Low-exposure countries should initially concentrate on building digital infrastructure to capitalize on AI complementarities as exposure grows.

Third, the sectoral exposure dimension contains much more information than the national one. Thus, policy actions need to be implemented at the sectoral level. Sectoral exposure varies, and even countries with low exposure have sectors with very high exposure levels. Take the case of India, which has relatively low aggregate exposure to AI among 133 countries, thanks to its large agricultural and informal services sectors, but ranks 36th out of 133 on the 2024 Oxford AI Readiness Index. Despite having low aggregate exposure, the country's formal IT services sector, which employs roughly 5.4 million individuals as programmers, call center operators, and analysts, is among the most exposed in the world. Aggregating exposure hides this bimodal distribution of sectoral exposure. Thus, the country's AI strategy should emphasize reskilling in the formal IT sector, leverage AI complementarities in higher education, and expand digital infrastructure in non-exposed sectors.

Fourth, the null result post-2022 should be interpreted with caution. Consistent null results across all six regression specifications indicate the absence of an impact on aggregate country-level data in the three years since LLM diffusion began, despite emerging firm- and task-level evidence of shifts in hiring patterns. Hence, governments should: (i) invest in more frequent data collection on the labor market, including administrative payroll data, online job vacancies, and quarterly task-based surveys; (ii) work with the relevant international bodies, such as the OECD and ILO, to create international data exchanges, so that changes in labor markets in small countries do not remain hidden; (iii) avoid passing any laws assuming the existence of labor displacement due to AI that are not backed by empirical facts. None of the above tasks would be difficult or expensive to undertake, as several of the required data series are already collected monthly through administrative social security and payroll systems.

Taken together, these implications point to three fiscal priorities for the next five years: (i) targeted reskilling programs according to sectoral exposure; (ii) labor adjustment policies such as wage insurance, portable benefits, and mobility grants to facilitate labor reallocation; and (iii) digital infrastructure expansion to increase the number of workers and firms benefiting from AI complementarities.

## 7. CONCLUSION

This paper estimates the effect of AI exposure on unemployment rates across 133 countries between 2010 and 2024 using a two-stage least squares approach with country- and year-fixed effects. In each of the main model specifications and robustness checks, AI exposure decreases unemployment in the diffusion period after 2018. This result holds for different AI exposure metrics (FRS and Eloundou, strongly correlated at 0.99), different instruments (Oxford and the IMF's AIPI), and excluding COVID-19 years. The post-2022 LLM-era coefficient for aggregate unemployment is statistically indistinguishable from zero across all six specifications, suggesting that the aggregate labor-market footprint of generative AI is not yet clear in the country-year dataset.

The paper makes three contributions. Methodologically, the paper develops a 133-country index of AI exposure, validated by two independent occupation-scoring algorithms, and identifies the LLM effect using two alternative instruments developed by independent organizations. Empirically, the paper finds a consistent negative relationship between AI exposure and unemployment in the pre-LLM diffusion period and an estimate statistically indistinguishable from zero in the LLM period, which holds under various reasonable measurement and identification alternatives. Substantively, the paper shows that aggregate unemployment rates across multiple countries are insufficient for studying the short-run effects of LLMs on employment, and that researchers and policymakers should look to more disaggregated data going forward.

Overall, the paper's results suggest that AI has been driving productivity and reallocating workers rather than displacing them so far. Nevertheless, this outcome cannot be regarded as a certainty: the 2022-and-later confidence intervals cover both positive and negative effects; the lack of an identifiable signal within three years does not mean there will be none within, say, five years. Therefore, while the threat of AI-driven labor-market disruption cannot be conclusively proven today, it is crucial for policymakers to build capacity to detect such displacement in advance and to implement complementary reskilling initiatives, benefit portability, and access to high-frequency microdata to respond swiftly in the event of actual displacement.

There are three key limitations to the paper's results. First, pre-trend placebo tests show that countries exposed to AI had already been experiencing different unemployment trends before 2018, meaning that some portion of the post-2018 coefficient captures pre-existing differences rather than reflecting AI's effect (Appendix A2). Second, the post-2022 null may reflect insufficient power due to only three post-LLM diffusion years. Finally, while this paper regards AI exposure as an invariant measure, it may change as AI progresses.

All three limitations point to specific methodological changes in future studies. First, quasi-experimental approaches — for example, difference-in-differences with multiple pre-treatment periods, or synthetic control methods — can help isolate AI effects from the pre-existing trends revealed by the placebo tests. Second, switching to high-frequency microdata that capture events before any observable aggregate effects, such as payroll and vacancy data, and to detailed survey data would make it easier to detect LLM effects at earlier stages. Third, updating the exposure score to account for evolving AI technologies would address the last limitation mentioned above. Finally, conducting sector-level analysis based on the bimodal exposure distributions observed in the paper (the India case) could further uncover hidden effects in aggregate data.

# REFERENCES

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. Journal of Labor Economics, 40(S1).

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. Journal of Political Economy, 128(6), 2188–2244.

Anthropic. (2026). Nowcasting the impact of AI on the labor market. Anthropic Economic Research.

Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States. American Economic Review, 103(6), 2121–2168.

Cazzaniga, M., et al. (2024). Gen-AI: Artificial intelligence and the future of work. IMF Staff Discussion Note SDN/2024/001.

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. arXiv:2303.10130.

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence. Strategic Management Journal, 42(12), 2195–2217.

Hardy, W., Keister, R., & Lewandowski, P. (2018). Educational upgrading, structural change and the task composition of jobs in Europe. Economics of Transition, 26(2), 201–231.

Oxford Insights. (2024). Government AI Readiness Index 2024.

Rajan, R. G., & Zingales, L. (1998). Financial dependence and growth. American Economic Review, 88(3), 559–586.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Identification and Inference for Econometric Models (pp. 80–108). Cambridge University Press.

# APPENDIX

## A0.1 Source Scoring Systems

We construct two country-level indices of AI exposure based on two independent occupational-level measures of AI exposure. One is the AI Occupational Exposure (AIOE) measure developed by Felten, Raj, and Seamans (2021). The AIOE maps the ten AI capabilities identified in the Electronic Frontier Foundation's AI Progress Measurement project into the fifty-two O*NET work abilities, and then combines these into a country-level measure based on the Standard Occupational Classification (SOC).

The FRS measure comprises 774 SOC occupations, is normalized to have a mean of 0 and a standard deviation of 1, and was constructed pre-LLM, prior to the publication of GPT-3.5 in late 2022. Thus, it reflects the capabilities of artificial intelligence available before the emergence of GPT-3.5 up until 2019–2020. The second measure we use is the Large Language Model (LLM) Exposure measure developed by Eloundou, Manning, Mishkin, and Rock (2023). It scores 19,265 O*NET tasks based on whether an LLM reduces the time required to complete the task by at least 50 percent without sacrificing quality (E1, full substitution) or by 25–50 percent (E2, partial augmentation). Scores of O*NET tasks are aggregated to the occupation level by the following rule:

$$\text{score}_o : \text{ELO} = E1_o + 0.5 \cdot E2_o$$

This implies that in cases where an LLM allows a reduction in task-completion time by half (full substitution), full weight is attributed to it, while in the case of partial substitution only partial weight. The resulting measure involves 923 O*NET-SOC occupations, takes values in the [0, 1] range, and is scaled in terms of task probabilities.

## A0.2 SOC-to-ISCO-08 Crosswalk

Occupational employment data are collected at the International Standard Classification of Occupations (ISCO-08) level, while both exposure measures are defined at the SOC classification level. To translate SOC-based occupational exposure scores into ISCO codes, we use the mapping table provided by Hardy, Keister, and Lewandowski (2018). It maps the American Occupational Classification System SOC-10 occupations to ISCO-08 unit groups. When several SOC codes correspond to a particular ISCO code, we take simple averages of the SOC-based scores for those cells. Aggregation to the ISCO-08 two-digit occupation category level yields 40 sub-major occupations under the FRS classification and 39 under the Eloundou classification.

## A0.3 Country-Level Aggregation

Occupation-level employment data are collected in the ILOSTAT database under the two-digit ISCO employment categories. We use occupation shares measured in 2010 as the pre-shock baseline and fix the occupational structure across all countries in advance. This means that variations in the index across countries reflect only differences in occupational structures in the period before the shock. The country-level AI exposure index is

$$\text{Exposure}_c = \sum_{o \in O} s_{c,o}^{2010} \cdot \text{score}_o$$

$s_{c,o}$ is country c's 2010 employment share in group o, and $\text{score}_o$ is that group's exposure score (corresponding FRS or Eloundou score). This shift-share construction is identical to that used by Cazzaniga et al. (2024) and structurally parallel to Autor, Dorn, and Hanson (2013).

## A0.4 Cross-Validation of the Two Indices

The two indices show a very strong correlation at 0.99 among the 133 countries considered in our sample estimation (Pearson r = 0.985, Spearman ρ = 0.986, Appendix A5 for the scatter plot). This strong correlation offers some interesting insights. Considering that the FRS and Eloundou systems were designed separately, by separate teams, for different AI generations, and using different mapping mechanisms, such a strong correlation suggests that country exposure depends more on the aggregation mechanism, namely occupational structure of the country rather than on any characteristic of the scores themselves. As a result, the findings do not depend on the choice between the FRS and Eloundou scores, and both are presented in Table 1.

## A0.5 Summary Statistics

Table A0 presents summary statistics for the merged panel.

**Table A0: Summary Statistics**

| Variable | Obs | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| Unemployment rate (%) | 2,020 | 7.361 | 5.504 | 0.119 | 36.472 |
| FRS exposure (rebuilt) | 2,085 | −0.357 | 0.305 | −1.211 | 0.445 |
| Eloundou exposure (rebuilt) | 2,085 | 0.258 | 0.056 | 0.123 | 0.415 |
| Oxford AI Readiness (0–1) | 2,085 | 0.488 | 0.168 | 0.169 | 0.870 |
| ln GDP per capita | 2,077 | 8.641 | 1.470 | 5.379 | 11.833 |
| ln Population | 2,085 | 15.930 | 1.941 | 10.533 | 21.095 |
| Services share (% GDP) | 2,056 | 56.106 | 10.945 | 6.448 | 94.148 |
| Industry share (% GDP) | 2,052 | 24.524 | 9.395 | 2.086 | 76.034 |
| Inflation (CPI, %) | 1,994 | 6.715 | 20.415 | −12.297 | 557.202 |
| GDP growth (%) | 2,081 | 3.271 | 4.945 | −32.909 | 63.335 |

*Notes: Summary statistics for the merged panel (139 countries with at least one non-missing observation, 2010–2024). Observation counts vary due to missing values across sources. FRS and Eloundou exposure indices are time-invariant* *country-level measures. Oxford AI Readiness is the 2024 cross-section normalized to the unit interval. The estimation sample (N = 1,937) is the intersection of non-missing unemployment, exposure, and Oxford scores.*

## A1. Identification Assumptions

The identification strategy relies on three key assumptions. First, the country-fixed effects $\mu_c$ include all factors that could affect exposure to artificial intelligence (AI), such as institutions, geography, legal origin, colonial history, and occupational structure. Second, year-fixed effects $\lambda_t$ absorb global shocks that affect all countries equally, including the pandemic-driven global recession of 2020–2021, global monetary arrangements, and movements in commodity prices. With these factors accounted for, $\beta_1$ and $\beta_2$ can be estimated by comparing the dynamics of within-country unemployment rates across countries with different exposure levels, measured in 2018 and 2022.

Finally, the instrument is designed so as to address potential problems associated with the endogeneity of the product $\text{Exposure}_c \times \text{LLM}_t$. Specifically, because the FRS occupational scores were developed before the publicly available versions of language models were launched, they cannot serve as a measure of exposure to LLMs post-2022, thereby generating measurement error that will likely make OLS estimates conservative. As the Oxford government readiness index 2024 relies on slow-moving factors that existed prior to the LLM shock, Oxford readiness is likely to be exogenous to the post-2022 unemployment rate once macroeconomic controls and fixed effects are controlled for. The identifying assumption is therefore that Oxford readiness affects unemployment only through exposure to AI. The support for this assumption comes from replacing the Oxford index with the IMF index, yielding estimates of the second-stage coefficients that are practically the same (columns A1 and B1 of Table 1), and from observing that the standardized post-2018 coefficients of both indexes perfectly align.

## A2. Pre-Trend Placebo Tests

A standard concern with shift-share exposure measures is that the cross-sectional component may be correlated with country trends in pre-treatment unemployment. We test this directly by regressing the change in average unemployment between 2014–2016 and 2017–2019 on the three predictors: FRS exposure, Eloundou exposure, and Oxford AI Readiness. All three result in negative coefficients significant at the 1% significance level (Table A1). Countries that later became more exposed to AI, or scored higher on Oxford readiness, were already experiencing decline in unemployment over the 2014–2019 pre-period.

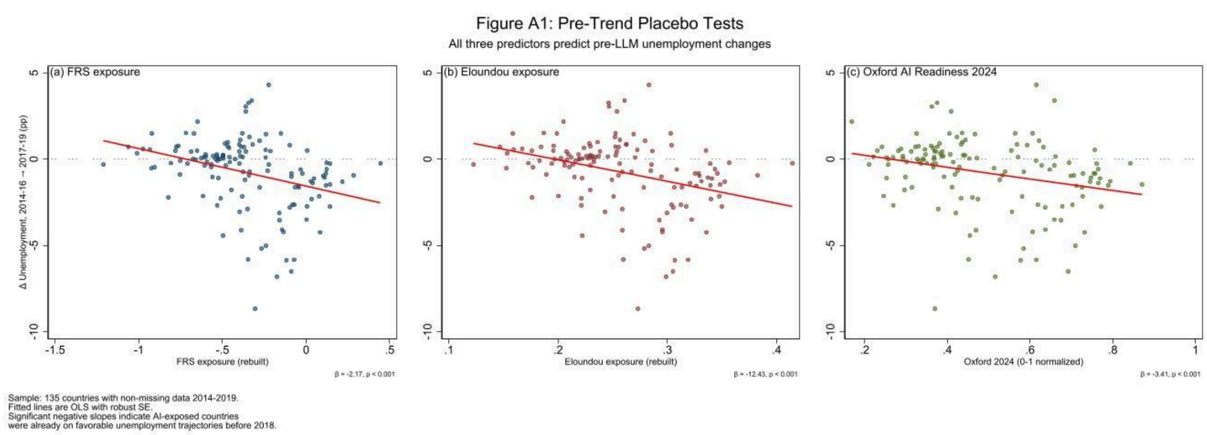


**Figure A1: Pre-Trend Placebo Scatters**

Two observations reduce the magnitude of the concern but do not completely eliminate it; excluding the COVID years (A3.3) yields a post-2018 coefficient implying a six-year cumulative decline of −2.6 percentage points, larger than a linear extrapolation of the 2014–2019 pre-trend over the same horizon. This points to the effect of AI exposure over and above the pre-existing trend. Additionally, the post-2022 LLM coefficient is statistically insignificant with exposure in every specification. If the pre-trend simply carried forward, we would expect the LLM coefficient to be at least as negative as the post-2018 coefficient; instead, it flips sign. We therefore read $\beta_1$ as a mix of pre-existing structural differences and a modest incremental AI-diffusion effect, and our design is unable to distinguish between the two.

**Table A1: Pre-Trend Placebo Regressions**

| Predictor of ΔU | Coeff. | Robust SE | p-value | $R^2$ | N |
|---|---|---|---|---|---|
| FRS exposure | −2.168*** | 0.389 | 0.000 | 0.104 | 135 |
| Eloundou exposure | −12.432*** | 2.302 | 0.000 | 0.113 | 135 |
| Oxford 2024 readiness | −3.408*** | 0.816 | 0.000 | 0.078 | 135 |

*Notes: OLS regressions of Δ average unemployment (2014–16 → 2017–19) on each predictor. Robust SE. All three predictors significantly predict changes in pre-LLM unemployment.*

## A3. Instrument Validity Checks

### A3.1 Weak-Instrument and Under-Identification Diagnostic Test

According to Kleibergen-Paap rk Wald F-statistics (see columns A1–B3 of Table 1), the F-statistic is large and varies between 106 and 191. At the 10% level, Stock-Yogo's critical value of the F-statistics is 16.38; hence, there are no problems with weak-instrument issues. Also, according to the Kleibergen-Paap LM statistic, we reject the null of no relevance at $p < 0.001$. Since the model is exactly identified (there is one instrument and one endogenous regressor), the Hansen J test is not applicable. Notably, the almost identical second-stage coefficients estimated by the two instruments (Oxford and IMF) indirectly support the exclusion restriction hypothesis.

### A3.2 Macroeconomic Controls

Macroeconomic controls align with the labor-market theory. Log population is positive (values around 14–15) in line with the empirical results. The log GDP per capita has the expected negative sign and is statistically significant at the 5% level in the full-sample regressions (A3 and B3), and marginally significant ($p \approx 0.05$) in the 122-country columns. The signs of the sectoral shares are positive (as expected), but they fail to reach statistical significance once fixed effects and log income are included in the regression. The GDP growth rate has the expected negative sign and is statistically significant at the 5% level in the full sample. Inflation is never statistically significant across any columns and is included only as a control in cross-country unemployment regressions. The inclusion of macroeconomic controls is robust to changing the exposure indicator and the instrument.

### A3.3 Robustness: Excluding the COVID Years (2020–2021)

It must be noted that the post-2018 window includes the 2020–2021 global economic crisis related to the COVID-19 pandemic, which increased unemployment due to associated lockdowns and government policies, rather than increased exposure. To assess the effects of the pandemic period specifically, we rerun the estimation, removing these years from the data. The post-2018 coefficient increases from −1.819 for FRS and −10.647 for Eloundou to −2.567 and −14.791, respectively. Both remain significant at the 1% level. The coefficient for LLM remains statistically insignificant and positive. The use of pandemic-year fixed effects attenuates $\beta_1$ in the full sample. The first-stage F-statistics drop slightly but stay well above the Stock-Yogo critical value of 16.38.

**Table A2: Robustness: Excluding COVID Years (2020–2021)**

| | (A3) Full | (A3) No-COVID | (B3) Full | (B3) No-COVID |
|---|---|---|---|---|
| AITrend ($\beta_1$) | −1.819*** | −2.567*** | −10.647*** | −14.791*** |
| | (0.590) | (0.765) | (3.306) | (4.257) |
| $\hat{Z}_{ct}$ ($\beta_2$) | 0.869 | 1.744 | 5.059 | 10.017 |
| | (0.895) | (1.356) | (4.856) | (7.378) |
| Observations | 1,937 | 1,681 | 1,937 | 1,681 |
| KP F-stat | 110.7 | 79.9 | 119.5 | 83.4 |

*Notes: IV (2SLS) estimates. No-COVID specifications drop 2020 and 2021. All columns include country- and year-fixed effects and the full macroeconomic control vector. Standard errors clustered at the country level. *** p < 0.01.*

### A3.4 Alternative Instrument: IMF AI Preparedness Index

We conduct the sensitivity test by replacing the Oxford Government AI Readiness Index 2024 with the IMF AI Preparedness Index (AIPI) 2023 (Cazzaniga et al., 2024). The index is developed independently by another organization based on four sub-indices: digital infrastructure, human capital and labor market policies, innovation and economic integration, and regulation and ethics. The index is scaled to take values between 0 and 1 and covers 174 countries; 122 countries overlap with the regression sample, adjusted for the control variables, unemployment, and exposure.

Table 1 presents the main estimation results based on the IMF index. We can see that the post-2018 coefficients of −1.608 (FRS) and −9.516 (Eloundou) differ from those based on Oxford only by about 8% and remain significant at 1% significance level. The post-2022 coefficients remain statistically insignificant. Importantly, the first-stage F-statistics are higher than those in Oxford-instrumented regressions (174.0 for FRS and 190.5 for Eloundou). Using independent instruments from different sources with almost identical coefficients constitutes strong proof of the validity of the exclusion restriction.

## A4. Country Sample and Exposure Ranking

The estimation sample comprises 133 countries across all income levels. The countries with the highest exposure include Singapore, Luxembourg, Switzerland, Israel, the Netherlands, the USA, Japan, Australia, and Germany — countries with a developed services sector employing a skilled workforce in professional, scientific, and technical fields. The less-exposed countries with substantial agricultural and informal sectors and poor formal service sectors are Burundi, Mozambique, Rwanda, Ethiopia, Madagascar, and Uganda. Note that India belongs to the bottom quartile of exposure, even with a very high readiness score. Hence, this country serves as a good case study example (Section 6).

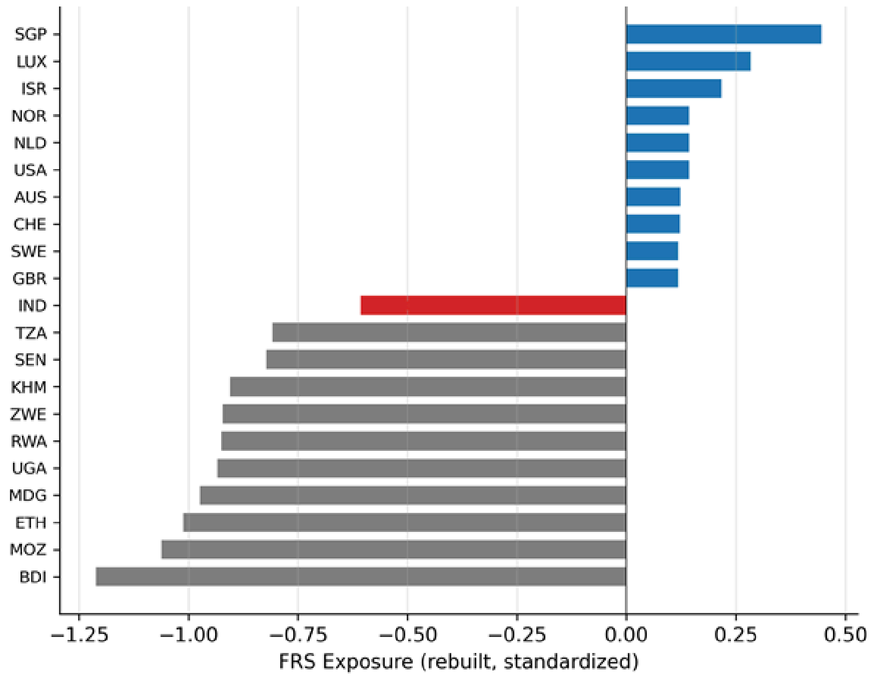


**Figure A2: Ranking of Countries According to AI Exposure (Top and Bottom 10) and India**

Figure A3 plots the two rebuilt exposure indices — the Eloundou-based index against the FRS-based index — across the 133 estimation countries, with the fitted line summarizing their near-linear relationship.

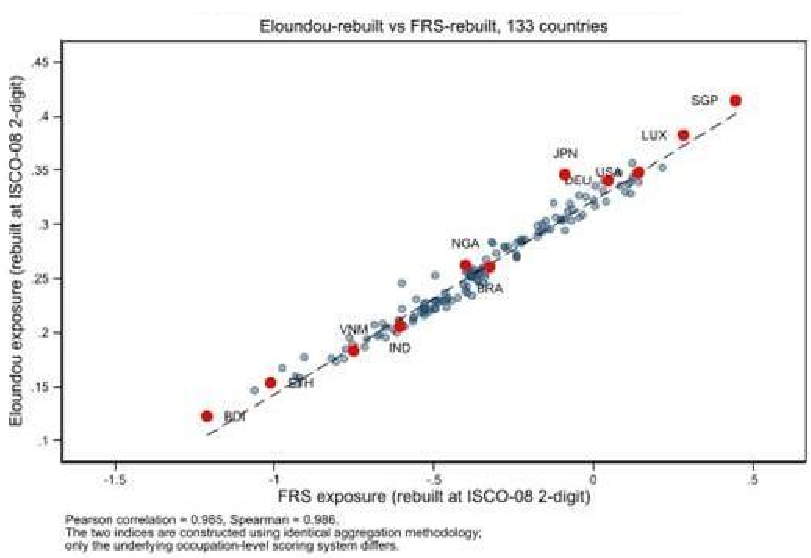


**Figure A3: Cross-Validation of AI Exposure Indices (Pearson = 0.985, Spearman = 0.986)**

The correlation between two exposure indicators among the 133 estimation countries is extremely high: Pearson $r = 0.985$; Spearman $\rho = 0.986$. Given that the two scores have been generated independently for different AI generations using different mappings of capabilities to occupational categories, it is clear that what drives the high exposure indicator is not a specific indicator but the national occupational structure itself, as reflected in the cross-walking procedure from SOC to ISCO-08. This justifies the claim in the text that identification is independent of the choice of scoring system.

## A5. Exposure-Readiness Quadrant

Figure A4 depicts the relationship between each country's AI exposure and Oxford 2024 readiness index; four particular quadrants stand out. Singapore and Luxembourg fall into the high-exposure, high-readiness quadrant. Burundi and Ethiopia represent the quadrant of low exposure and low readiness. The US, Germany, and Japan also sit in this high-exposure, high-readiness region. Finally, India, which serves as the case study in Section 6, falls in the low exposure and high readiness quadrant; this bimodal pattern across countries explains why aggregated country-level indices may miss relevant policy implications.

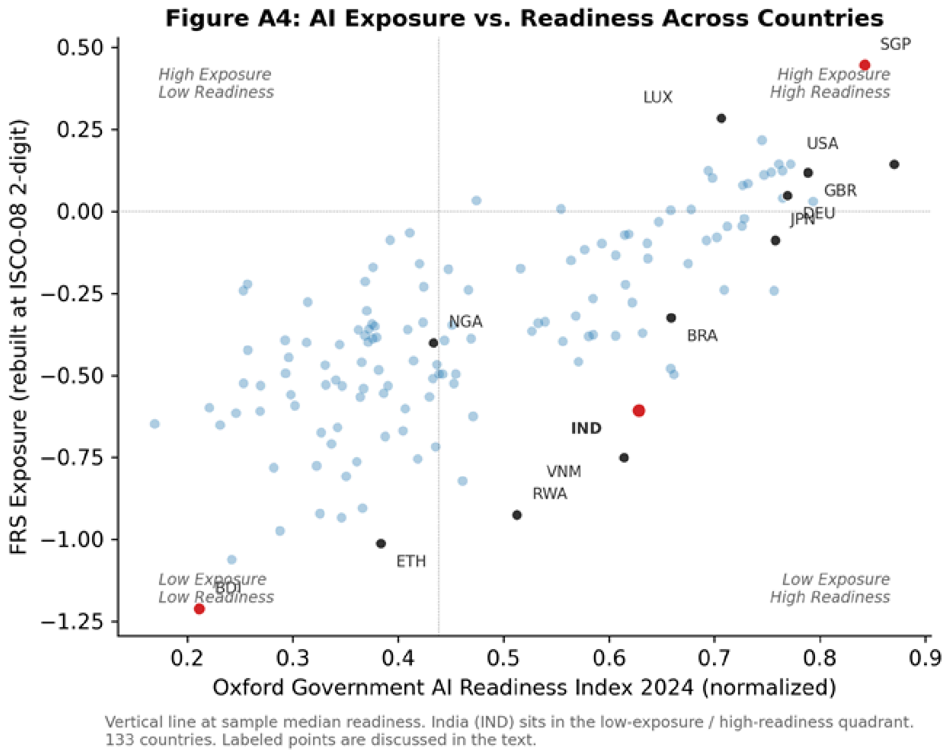


**Figure A4: AI Exposure vs. Readiness Across 133 Countries**

Taken together, the quadrants confirm that exposure and readiness are only weakly aligned across countries: high-readiness economies such as India can still sit in the low-exposure region, which is precisely the pattern that motivates the sector-level policy focus in Section 6.

Contact Us:

https://icpp.ashoka.edu.in
icpp@ashoka.edu.in

Scan the QR code to follow us on LinkedIn

**Delhi Office:**
74, Okhla Phase III,
Okhla Industrial Estate,
New Delhi, Delhi 110020

**Campus Address:**
Plot No. 2, Rajiv Gandhi Education City,
National Capital Region,
P.O. Rai, Sonepat, Haryana - 131029 (India)