ISSN: 2392- 5418 **e-ISSN:** 2520- 7423 **DOI:**10.59799/SQZH5356

Artificial Intelligence Adoption and Labor Productivity in Saudi Arabia: Econometric Evidence within the Framework of Saudi Vision 2030

Original Research Article

Amani Mohammed Aldukhail

Department of Economics,

College of Business and Economics, Qassim University, Qassim, Buraydah 52571, Saudi Arabia

Email: aldukhail@qu.edu.sa

ISSN: 2392- 5418 **e-ISSN:** 2520- 7423 **DOI:**10.59799/SQZH5356

Artificial Intelligence Adoption and Labor Productivity in Saudi Arabia: Econometric Evidence within the Framework of Saudi Vision 2030

Abstract

Objective: This study investigates whether the expansion of national capacity for AI adoption was associated with higher labor productivity and whether the relationship differed between the short run and the long run in Saudi Arabia, 2005–2024. It also tests whether the findings are robust to varying index constructions and structural controls.

Material and Methods: The annual macroeconomic data were analyzed using an autoregressive distributed lag (ARDL) framework. Labor productivity is defined as gross domestic product per person employed in constant 2021 purchasing-power-parity dollars. Following systematic screening for conceptual relevance and time-series coverage, an AIAI was formed aggregating Internet usage, fixed-broadband subscriptions and tertiary enrollment. Further, gross fixed capital formation and trade openness were included as control variables. We performed unit-root tests and a bootstrap bounds test, estimated an equilibrium-correction model, and conducted residual and stability diagnostics as well as robustness specifications using principal-component weights, leave-one-component-out indices, structural-period indicators and oil rents.

Results: While AIAI grew from 0.005 in 2005 to 1.000 in 2024, labor productivity decreased from about USD 182,020 to USD 131,865 per employed person. Each of the baseline variables was integrated of order one. Cointegration was supported by the bootstrap bounds test for cointegration (F = 16.036; upper-tail p = 0.00095). Both the long-run AIAI coefficient (beta = -0.0909; 95% confidence interval, -0.1613 to -0.0205; p = 0.01141) and short-run change coefficient (beta = −0.6815; p = 0.00031) were negative and statistically significant. We observed an equilibrium-correction coefficient of -1.1109 (p = 0.0000001). The negative sign remained in all robustness specifications, but precision was sensitive to index construction and the short annual sample.

Conclusions: Measurable aggregate labor-productivity gains through AI-enabling capacity were not realized in Saudi Arabia during 2005–2024. This finding is in line with transition lags, incomplete diffusion, differential measurement of readiness and actual use, and the need for complementary skills, data, organizational redesigns, and sector-level implementation. This should not be taken as evidence that AI causes a decline in productivity.

Keywords: artificial intelligence; labor productivity; ARDL; digital transformation; Saudi Vision 2030

1. Introduction

AI is increasingly viewed as a general-purpose technology that potentially rewrites how firms combine labor, capital, data, and knowledge. What economic value it has comes from its potential for human-in-the-loop automation, judgment augmentation, better prediction, faster

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

knowledge production, and new product and process creation.[1,2] Generative AI has decreased time to completion and increased the quality of writing performance in professional domains.[3] Additionally, it has increased the productivity of customer-support agents (with the biggest gains among less experienced workers),[4] and AI investments have been linked to firm growth and product innovation.[5] Similar opportunities for productivity are emerging in other areas of research-based knowledge work.[6]

This microeconomic gain need not show up straight away in national productivity numbers. Such general-purpose technologies need to be supported by a complementary investment in digital infrastructure, data governance, skills and organizational capital, cybersecurity and a redesign of processes. Measured productivity may have a J-curve shape during adjustment as the costs of intangible investment are incurred before their output benefits manifest.[7] Thus, macroeconomic estimates will depend on task exposure and the speed of adoption, sectoral composition, market structure, and how well institutions and workers can leverage new tools.[8-12] The difference between technical capability and economically effective diffusion is key to aggregate assessment.

Early AI adoption is more concentrated among large, digital and human-capital-intensive firms,[13] while cross-country microdata reveal productivity premia reflecting complementary human and technological capital.[14,15] These patterns thus create an aggregation problem: one can have strong progress among leading adopters, and at the same time weak national productivity growth when smaller firms and less digital sectors adopt slowly. Gaps in connectivity, compute, context-specific data, skills and institutional quality can prolong the realization of benefits from AI for developing and transforming economies.[16]

Saudi Arabia provides a policy-relevant environment within which to investigate this disconnection. Digital transformation — and innovation and human-capital development — along with economic diversification, is central to the Kingdom's national model of development, Saudi Vision 2030.[17] AI capabilities, AI talent development, data governance and public-sector adoption of AI have all been expanded through such initiatives as the National Strategy for Data and AI.[18,19] Digital-government measurement also tracks substantial institutional progress.[20] These developments coincided with labor-market reform, investment expansion and sectoral diversification,[21,22] so that direct inference of productivity effects solely from readiness indicators seems limited. Forward-looking simulations project higher Saudi output under sustained AI-market growth,[23] though ex-post econometric evidence remains thin on the ground.

The majority of Saudi studies focus on digitalization, e-government, sustainable growth or sector-specific applications rather than the aggregate relationship between AI-adoption capacity and output per worker. This study raises the question of whether the expansion in AI-enabling capacity was associated with higher Saudi labor productivity over the period 2005–

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

2024. It uses capital formation and trade openness as control variables to evaluate the positive conditional link; it distinguishes short-run dynamics from long-run equilibrium; and tests robustness with respect to different index designs and structural controls. This is a cautious historical estimate that separates readiness, direct adoption and realized productivity under Vision 2030.

2. Material and Methods

2.1. Study Design, Scope, and Data Protocol

The study adopts a quantitative longitudinal econometric framework covering 20 annual observations for Saudi Arabia from 2005 to 2024. The unit of analysis is the country-year. The dependent variable is labor productivity (LP), as measured by the World Development Indicators (WDI) series GDP per person employed in constant 2021 purchasing-power-parity dollars (series SL.GDP.PCAP.EM.KD).[24] The natural logarithm of LP is used to reduce scale dependence and allow approximate percentage interpretation of coefficients. Gross fixed capital formation (GFCF; NE.GDI.FTOT.ZS) and trade openness (OPEN; NE.TRD.GNFS.ZS) are expressed as percentages of GDP and act as macroeconomic controls. On 15 August 2026, data were extracted from internationally harmonized WDI records.[25]

The aggregate AI-use statistics at the level of Saudi establishments are sparse prior to recent years and hence cannot be employed in a 20-year time-series model. The General Authority for Statistics reports the methodology for the establishment ICT survey and a recent AI-use rate of 33.1% of establishments.[26,27] Those observations are taken to serve as an external-validity benchmark rather than being woven into the historical estimation series. It was a deliberate choice not to combine an adoption-specific measure with earlier readiness proxies that differ in both their definitions and sampling frames.

2.2. Variable Measurement and AI Adoption Index

AI Adoption Index (AIAI) is an index that helps us assess the national capacity for AI diffusion, not direct adoption. Indicators were chosen based on three criteria: theoretical relevance to adoption capacity, annual comparability, and no more than two consecutive missing observations between 2005–2024. Data on Internet use, fixed-broadband subscriptions, secure Internet servers, tertiary enrollment and resident patent applications formed the basis of the first screen. This is consistent with multidimensional preparedness frameworks which delineate infrastructure, human capital, innovation capacity, and institutional conditions.[28] The secure-server and patent series were excluded from the baseline because their gaps exceeded the coverage rule, although patent evidence was used to inform construct screening.[29]

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

The included variables were: Internet users (% of population), fixed-broadband subscriptions per 100 people and gross tertiary enrollment. The first step was that each component was given a positive orientation and normalized over the sample to the interval 0–1. For 2023, a single missing tertiary-enrollment value was linearly interpolated between those for the years 2022 and 2024. Broadband observations for 2013–14 were not imputed; the index in these two years was based on the available normalized components. Equal weights define the baseline. In the situation where a component was missing, available component weights were rescaled to sum to one in the principal-component robustness model. Definitions and expected signs are presented in Table 1, and Figure 1 summarises the whole analytical process.

Table 1. Variables, Operationalization, Expected Signs, and Data Sources

Variable	Operational Measure	Transformation	Role	Expected Sign	Source
LP	GDP per person employed, constant 2021 PPP dollars	Natural logarithm	Dependent	-	World Bank WDI
AIAI	0-1 composite of Internet use, fixed broadband, and tertiary enrollment	Min-max; equal weight	Main explanatory	+	World Bank / ITU / UNESCO
GFCF	Gross fixed capital formation (% of GDP)	Level	Control	+	World Bank WDI
OPEN	Trade (% of GDP)	Level	Control	+/-	World Bank WDI
OIL	Oil rents (% of GDP), 2005-2021 robustness sample	Level	Robustness control	+/-	World Bank WDI

Note: LP = labor productivity; PPP = purchasing-power parity; AIAI = AI Adoption Index; GFCF = gross fixed capital formation; OPEN = trade openness; OIL = oil rents; WDI = World Development Indicators; ITU = International Telecommunication Union; UNESCO = United Nations Educational, Scientific and Cultural Organization.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

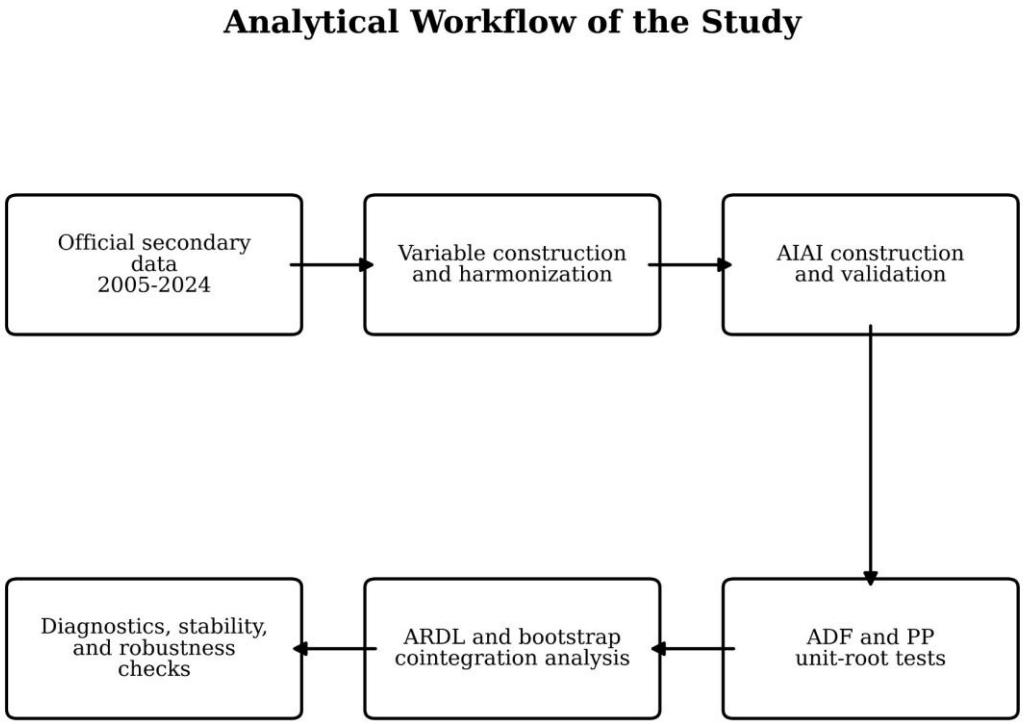


Figure 1: Analytical workflow from official secondary data and variable construction through index validation, unit-root testing, ARDL/bootstrap cointegration estimation, diagnostics, stability assessment, and robustness checks.

For component i in year t, normalization and aggregation were calculated as follows:

$$z(i,t) = [x(i,t) - \min x(i)] / [\max x(i) - \min x(i)]; \quad AIAI(t) = [1 / K(t)] \Sigma_i z(i,t)$$

where K(t) is the number of available components in year t. This formulation keeps the published index scale transparent, avoids arbitrary units, and permits direct inspection of how changes in national enabling capacity relate to productivity.

2.3. Econometric Strategy

The baseline long-run specification relates log labor productivity to the AIAI, GFCF, and OPEN:

$$\ln[LP(t)] = \beta_0 + \beta_1 AIAI(t) + \beta_2 GFCF(t) + \beta_3 OPEN(t) + \varepsilon(t)$$

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

The expected sign of AIAI is positive. GFCF is expected to underpin productive capacity (large transformation projects may, however, take time to pay off), while the expected sign of OPEN is theoretically ambiguous since openness can widen access to technology and competition but also reflects import dependence. Oil rents are not included in the baseline, as the available series ends in 2021; they thus enter a restricted-sample robustness model.

Specifically, augmented Dickey-Fuller and Phillips-Perron tests were performed first to establish that the time series are not integrated of order two. We chose the autoregressive distributed lag approach because it allows for I(0) and I(1) regressors, separates short-run adjustment from long-run association between variables, and is feasible with a small annual sample when lag complexity is restricted.[30] A maximum lag of one year was imposed, and selection was guided by the Bayesian information criterion to protect degrees of freedom. We tested cointegration by means of both the augmented and bootstrap bounds procedures, which were developed to improve small-sample inference.[31,32] Long-run coefficients and the equilibrium-correction representation were estimated using established ARDL implementation guidance.[33]

Inference outputs effect estimates, robust standard errors, confidence intervals and exact p-values (if the underlying output is capable of providing these results). The diagnostics included Breusch-Godfrey serial correlation, Breusch-Pagan unconditional heteroskedasticity, ARCH effects, Jarque-Bera residual normality test, Ramsey RESET functional form and CUSUM for parameter stability. HC3 standard errors were retained as the ARCH-LM test suggested conditional heteroskedasticity. All data manipulation, estimation, diagnostic checks and visualizations were performed using Python 3.12 with pandas, NumPy, statsmodels-compatible econometric routines and custom bootstrap code. Because an observational design does not fully identify causal effects, coefficients are interpreted as conditional associations.

2.4. Robustness, Ethics, and Reproducibility

For robustness analysis, one modeling decision was changed at a time. Principal-component weights, three leave-one-component-out indices, a Vision 2030 indicator, a COVID-19 indicator, both structural indicators jointly, and a 2005–2021 model that added oil rents were used. Sign stability across specifications was handled as evidence of directional robustness; p-value changes were interpreted in terms of precision, not as a binary validity test. Institutional review-board approval and informed consent were not required because the study was based on publicly available aggregate national statistics containing no individual-level or identifiable human-participant information. This report provides series identifiers, transformation rules, equations and diagnostic criteria as well as robustness definitions to facilitate independent replication.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

3. Results

3.1. Descriptive and Stationarity Results

For the baseline model, the sample includes 20 full annual observations. The AIAI grew from 0.005 in 2005 to 1.000 in 2024, documenting widespread growth in Internet access, fixed broadband and tertiary enrollment. Over the same period, real labor productivity decreased from approximately USD 182,020 to USD 131,865 per employed person. AIAI was contemporaneously correlated with log labor productivity (−0.856). In Figure 2, we plot this descriptive contrast between the two series; but it does not in itself demonstrate a causal relationship; rather, it justifies formal dynamic testing because both series exhibit strong time trends and structural changes that may be common to each.



Figure 2: AI Adoption Index and labor productivity in Saudi Arabia, 2005-2024. Labor productivity is min-max scaled only for visual comparability; the original constant 2021 PPP-dollar series is used in the econometric analysis. N = 20 annual observations.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

Descriptive statistics and unit-root test results are given in our Table 2. In levels, the Augmented Dickey–Fuller and Phillips–Perron tests do not reject the unit-root null. Once we first differenced, the tests rejected the null for all baseline variables (the exact p-values range from 0.00001 to 0.01465). Accordingly, these variables were all interpreted as I(1), with no evidence of I(2) behavior. This fulfills the integration condition of the ARDL bounds methodology.

Table 2. Descriptive Statistics and Unit-Root Tests

Variable	Mean	SD	Min	Max	ADF Level	PP Level	ADF First Difference	PP First Difference	Order
lnLP	11.914	0.093	11.790	12.112	-1.902 (0.33101)	-2.266 (0.18304)	-5.138 (0.00001)	-4.518 (0.00018)	I(1)
AIAI	0.502	0.342	0.005	1.000	-0.227 (0.93523)	-0.167 (0.94234)	-3.359 (0.01245)	-3.305 (0.01465)	I(1)
GFCF	24.152	2.762	19.312	30.630	-1.669 (0.44712)	-1.078 (0.72360)	-3.682 (0.00437)	-3.695 (0.00418)	I(1)
OPEN	71.126	15.717	47.533	96.103	-0.591 (0.87297)	-0.559 (0.87999)	-3.505 (0.00786)	-3.642 (0.00501)	I(1)

Note: N = 20. Values in parentheses are exact p-values. ADF = Augmented Dickey-Fuller; PP = Phillips-Perron; SD = standard deviation; I(1) = integrated of order one. First-difference tests use the change in each variable.

3.2. Cointegration and ARDL Estimates

The bootstrap bounds statistic was 16.036, surpassing the 99.9% upper critical value of 15.502; the empirical upper-tail p-value was 0.00095. Hence, the null of no level relationship was rejected. For the long-run, the AIAI coefficient was negative and statistically significant (beta = −0.0909; robust SE = 0.0359; 95% CI: −0.1613 to −0.0205; p = 0.01141). Because the dependent variable is logged and the index is measured from 0 to 1, a 0.10 increase in the AIAI corresponds to an estimated long-run difference of approximately -0.91% in labor productivity, conditional on the controls. This is a semi-elasticity interpretation, not a causal effect.

GFCF was also negative (beta = -0.0103; p = 0.00006), while trade openness was positive (beta = 0.0018; p = 0.03460). In the immediate term, the AIAI change coefficient was -0.6815

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

(p = 0.00031), GFCF was -0.0060 (p = 0.00834), and OPEN was 0.0071 (p = 0.00182). The error-correction coefficient was −1.1109 (p = 0.0000001), which means that slightly more than the previous year's disequilibrium was corrected within one year. It indicates convergence with overshooting rather than monotonic adjustment.

It gave R-squared = 0.979 and adjusted R-squared = 0.966 for the model. Serial correlation, unconditional heteroskedasticity, residual non-normality, functional-form error at the 5% level and parameter instability were not detected. However, the ARCH-LM test detected conditional heteroskedasticity (p = 0.00045), supporting HC3 inference. As shown in Figure 3, the CUSUM path did not cross the 95% boundaries. The full estimates and diagnostic decisions can be seen in Table 3.

Table 3. Bootstrap Bounds Test, ARDL Estimates, Model Fit, and Diagnostics

Panel / Parameter	Estimate	Robust SE	Statistic	p-Value	95% CI / Decision
Panel A: Bootstrap Bounds Test					
F-statistic	16.036	-	99.9% I(1) upper = 15.502	0.00095	Cointegration supported
Panel B: Long-Run and Short-Run Estimates					
Long run: AIAI	-0.0909	0.0359	z = -2.530	0.01141	[-0.1613, -0.0205]
Long run: GFCF	-0.0103	0.0026	z = -3.999	0.00006	[-0.0154, -0.0053]
Long run: OPEN	0.0018	0.0008	z = 2.113	0.03460	[0.0001, 0.0035]
Short run: change in AIAI	-0.6815	0.1320	t = -5.161	0.00031	[-0.9721, -0.3908]
Short run: change in GFCF	-0.0060	0.0019	t = -3.208	0.00834	[-0.0100, -0.0019]
Short run: change in OPEN	0.0071	0.0017	t = 4.080	0.00182	[0.0033, 0.0109]
Equilibrium-correction term	-1.1109	0.0925	t = -12.005	0.0000001	[-1.3145, -0.9072]
Panel C: Model Fit and Diagnostics					
R-squared	0.979	-	-	-	Model fit
Adjusted R-squared	0.966	-	-	-	Model fit

ISSN: 2392- 5418 **e-ISSN:** 2520- 7423 **DOI:**10.59799/SQZH5356

Panel / Parameter	Estimate	Robust SE	Statistic	p-Value	95% CI / Decision
Breusch-Godfrey LM	0.252	-	-	0.616	No serial correlation
Breusch-Pagan	2.806	-	-	0.902	No unconditional heteroskedasticity
ARCH-LM	12.311	-	-	0.00045	ARCH effects detected
Jarque-Bera	0.207	-	-	0.902	Normality not rejected
Ramsey RESET	3.969	-	-	0.074	No rejection at 5%
CUSUM	0.370	-	-	0.999	Stability not rejected

Note: N = 20. Long-run p-values use the reported z statistics; short-run p-values use the reported t statistics. HC3 robust standard errors are shown because ARCH effects were detected. CI = confidence interval; CUSUM = cumulative sum stability test.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

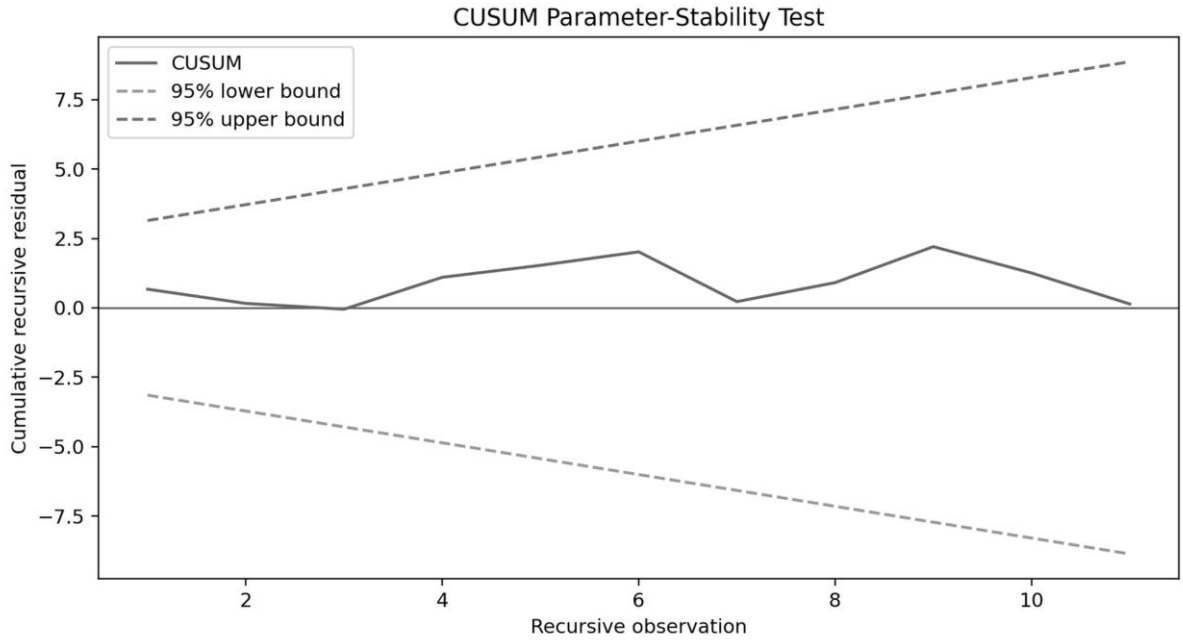


Figure 3: CUSUM parameter-stability test. The solid line is the cumulative recursive-residual path, and the dashed lines are the 95% lower and upper stability boundaries. The path remains within both boundaries.

3.3. Robustness Results

Table 4. Robustness of the Long-Run AI Adoption Index Coefficient

Alternative Specification	N	Long-Run AIAI Coefficient	Robust SE	Approximate 95% CI	p-Value
PCA-weighted AIAI	20	-0.1258	0.0645	[-0.2522, 0.0006]	0.051
Leave out Internet use	20	-0.0343	0.0881	[-0.2070, 0.1384]	0.697
Leave out fixed broadband	20	-0.0956	0.0885	[-0.2691, 0.0779]	0.280

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

Alternative Specification	N	Long-Run AIAI Coefficient	Robust SE	Approximate 95% CI	p-Value
Leave out tertiary enrollment	20	-0.1073	0.0804	[-0.2649, 0.0503]	0.182
Add Vision 2030 indicator	20	-0.1002	0.0541	[-0.2062, 0.0058]	0.064
Add COVID-19 indicator	20	-0.1217	0.0485	[-0.2168, -0.0266]	0.012
Add Vision 2030 + COVID-19	20	-0.1537	0.0299	[-0.2123, -0.0951]	0.0000003
Add oil rents, 2005-2021	17	-0.0886	0.0925	[-0.2699, 0.0927]	0.338

Note: AIAI = AI Adoption Index; PCA = principal component analysis; SE = standard error; CI = confidence interval. Approximate intervals equal estimate +/- 1.96 x robust SE. The exact two-sided p-value for the joint-indicator model was calculated from its reported robust z statistic.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

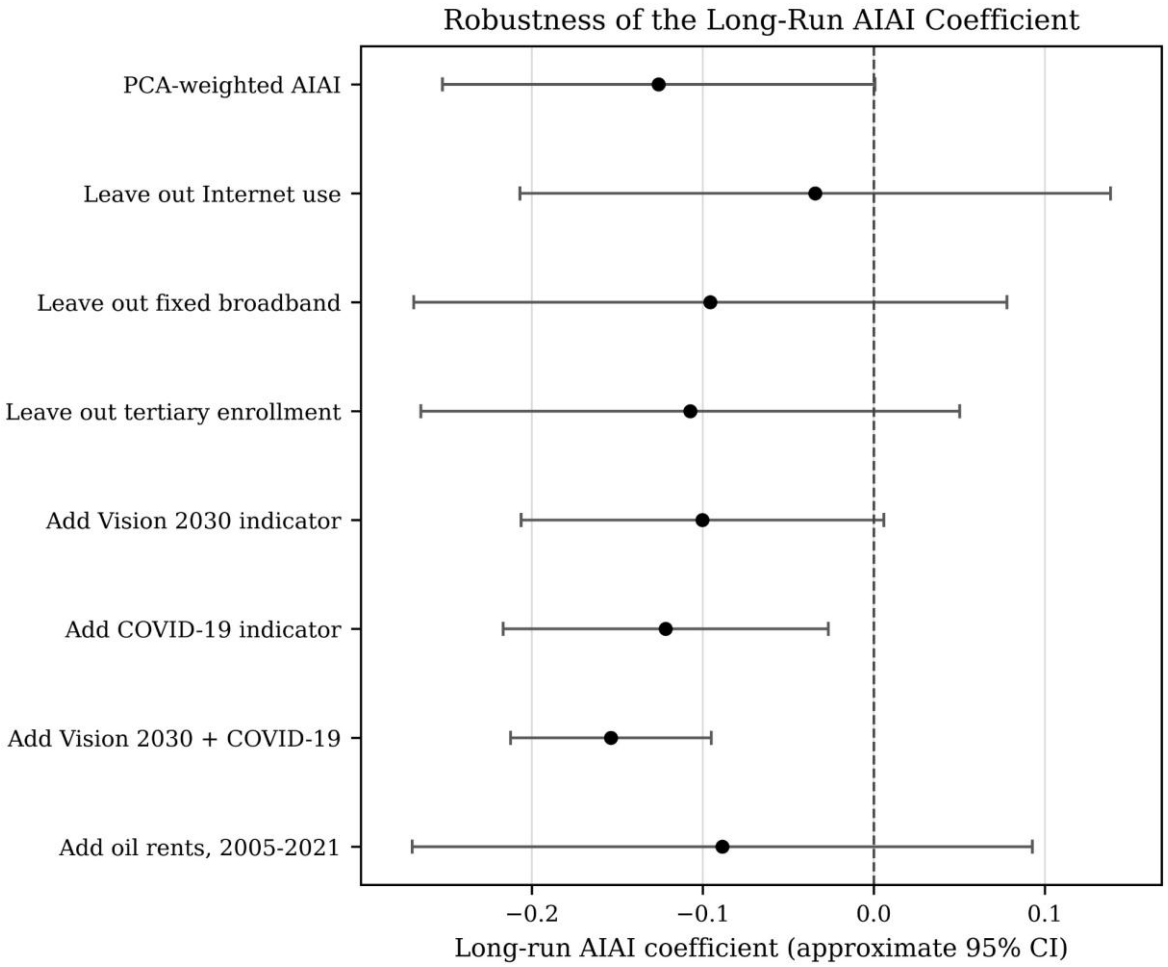


Figure 4: Robustness of the long-run AI Adoption Index coefficient. Points show coefficient estimates and horizontal bars show approximate 95% confidence intervals calculated as estimate +/- 1.96 robust standard errors. The vertical line marks zero. N = 20 except for the oil-rents specification (N = 17).

The long-run AIAI coefficient remained negative in each of eight alternative specifications (Table 4 and Figure 4). With principal-component weights, it was marginally significant (beta = -0.1258; p = 0.051); it was imprecise in the three leave-one-component-out models and negative but imprecise when a Vision 2030 indicator or oil rents were added. The COVID-19 indicator resulted in a significant coefficient of -0.1217 (p = 0.012). The joint Vision 2030 and COVID-19 specification had a slightly larger negative coefficient (-0.1537; p = 0.0000003). This exact two-sided value was computed from the reported coefficient and robust standard error based on the normal approximation applied to the remaining long-run robustness tests.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

These checks are consistent with a stable negative direction but not with stable statistical precision. The loss of significance apparent with component deletion is crucial: it indicates that the baseline association is tied to the combination of proxies in the readiness construct rather than to one dominant proxy. With a small sample, adding controls or structural indicators increases uncertainty as well. Thus, the baseline model does not support the positive study hypothesis but instead is more consistent with a historically negative association subject to measurement sensitivity than with evidence of an adverse causal effect.

4. Discussion

The main result is that the large jump in Saudi AI-enabling capacity was associated with lower instead of higher aggregate labor productivity — after conditioning on capital formation and trade openness over 2005–2024. This describes a far different experience than that which has been described in studies with employees and experiments at the workplace.[3-6] This difference is not a contradiction in itself. These studies observe direct AI use in narrow sets of tasks, whereas AIAI is a national-level stock of conditions that facilitate the use of AI. Connectivity and education can develop long before firms use AI at scale, or employees reorganize work around it, or the resulting value added becomes visible in output per worker.

This is consistent with evidence of the productivity J-curve and diffusion.[7,13-16] The early adoption costs entail investments in cloud and compute capacity, data cleaning and governance, cybersecurity, training and experimentation, and redesign of workflows. This outlay largely gives rise to intangible capital which finds only a weak reflection in current output. Alternatively, adoption could be heavily concentrated in government, large firms or service functions where the productivity gains do not immediately dominate economy-wide productivity. This transmission mechanism is summarized in Figure 5: productive AI use can only scale if readiness enables the combination of complementary assets and organizational change to turn technical access into recurring, scalable production improvements.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

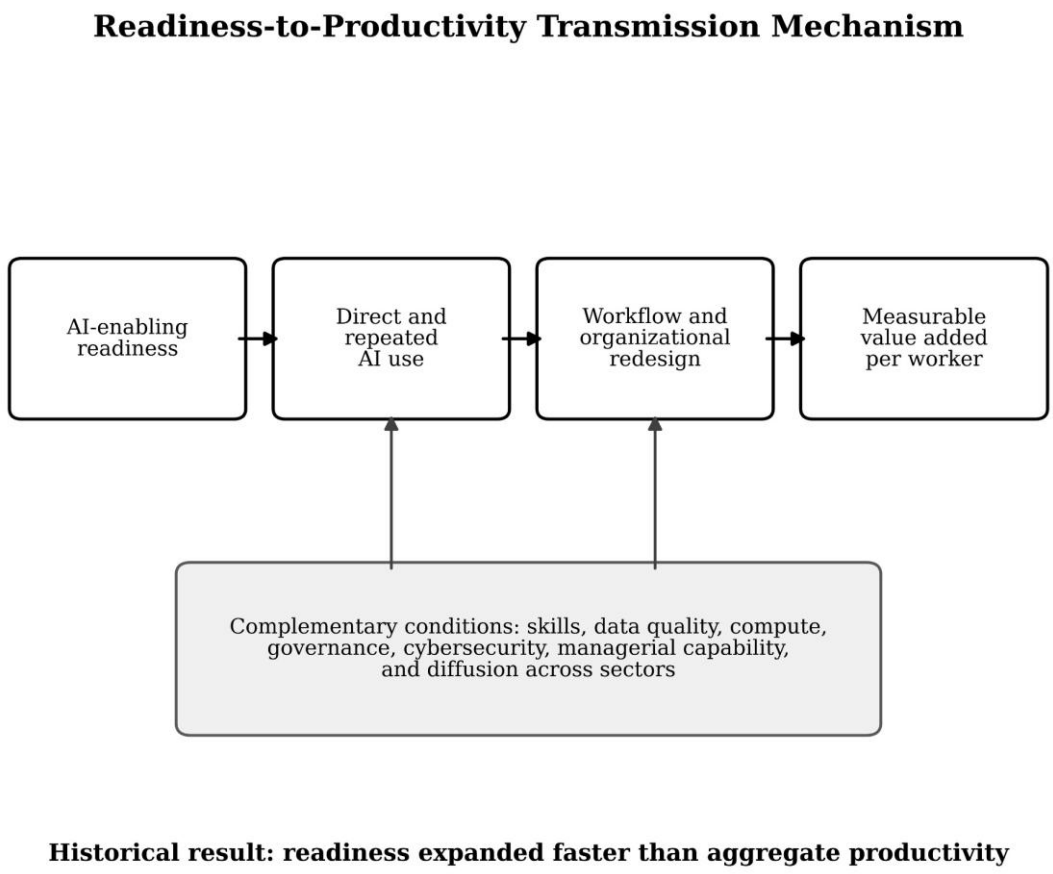


Figure 5: Conceptual readiness-to-productivity transmission mechanism used to interpret the results. Arrows represent a theorized sequence rather than separately estimated causal paths; complementary skills, data, compute, governance, and organizational redesign determine whether readiness becomes productive AI use and measurable value added per worker.

The timing is particularly relevant given the larger transformation of Saudi Arabia. At the same time, Vision 2030 quickened investment, modernized the public sector, and reformed labor markets while also promoting diversification.[17-22] These transformations can affect both the numerator and denominator of output per worker. Prior to being fully operative, new infrastructure and strategic projects could absorb labor; increases in labor-force participation and sectoral reallocation can also depress average measured productivity temporarily, even if long-run productive capacity is improving. The negative GFCF coefficient is thus a possible transition-period relationship, but its generalizability beyond the sample should not be taken for granted.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

In contrast, trade openness was positively related to productivity. Openness can also facilitate access to imported capital goods, software, intermediate inputs, managerial know-how and competitive discipline. Yet in the Gulf, there has been progress in government effectiveness, financial inclusion, resilience and growth, while persistent gaps in firm adoption and digital skills remain.[34] Our point estimate for OPEN is positive, as would be consistent with a channel where knowledge is transferred via trade, but cannot be interpreted as offering structural identification of the causal impact of trade policy.

The results also distinguish between estimates of historical data and projections for the future. The Saudi simulation cited above assumes that the expansion of AI markets will ultimately lead to enhanced sectoral efficiency and gross domestic product.[23] New international evidence similarly predicts aggregate positive gains that may be sizable when adoption is more widespread across sectors.[12] The present study, instead, asks whether enabling capacity and realized productivity moved together over a past transition period. In such cases, subsequent progress can therefore sit alongside a weak or negative historical association. In a similar vein, the World Bank's recent development framework also underlines that adoption dividends depend on electricity, connectivity, skills, context-specific data and institutional quality, rather than access alone.[16]

The second issue of interpretation addresses measurement. While the AIAI is transparent and reproducible, Internet use, broadband and tertiary enrollment are broad prerequisites, not observations of AI-intensive production. And that brings us to why direct statistics are important: the 33.1% estimate of establishment use.[27] National planning for statistical development ought to demarcate experimentation from deployment and register expenditures, utilization frequency, impacted tasks, worker exposure, complementary training provided, as well as realized output. A good organizing principle for such measurement is the four-Cs framework of connectivity, compute, context and competency.[35]

The study has five limitations. The proxy measures capacity rather than the intensity or quality of AI use. Second, lag length, statistical power and the number of controls are limited by 20 annual observations. Third, reverse causality or omitted structural factors are not eliminated by the ARDL design. Fourth, aggregate data hide significant heterogeneity at the level of industry, occupation, region and firm size. Fifth, the precision of coefficients is affected by index weights and missing-data rules, as the robustness results confirm. Such limitations refine the claim: the historical relationship identified is conditional, not an unambiguous measure of AI's direct impact on productivity.

Within these limits, the evidence adds three parts to the puzzle. What is new here is that it sheds light on a transparent long-run Saudi indicator as opposed to extrapolating from recent adoption surveys; treats short-run adjustment separately from long-run association; and finally reports directional and measurement robustness rather than reporting a single specification.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

The policy implication is not to slow AI investment but to manage the transition from readiness to productive use. Such diffusion may need to be aligned with initiatives at the sector or firm level, and therefore likely focused on managerial and worker training, interoperable data, secure compute access, process redesign, and evaluation based on value added per worker.

5. Conclusion

While the Saudi economy witnessed a radical scaling up of AI-enabling structures between 2005 and 2024, evidence from the econometric analysis suggests that many of these changes had not yet been manifested in higher aggregate labor productivity. In the baseline ARDL model, a negative long-run association is identified, and the sign continues to be negative in each robustness specification. This result is best understood as a transition and measurement finding. The four stages of diffusion exhibit a propensity to advance at different paces from one another or—equally significant—from organizational transformation and realized productivity.

So policy should thus focus less on counting digital capacity or how many AI initiatives a country has and more about productive use. Saudi institutions might be able to annually generate figures – specific to sectors and firms – on the intensity of adoption, tasks affected, capital expenditure on AI-related investments and data/compute use, worker training, and value added per employee. When selecting AI tools, firms should consider process redesign, data quality and cybersecurity, as well as workforce reskilling. Subsequent research could use longer direct-adoption series, linked firm panels, sectoral production data and causal or quasi-experimental designs to identify when and where AI investment begins to generate measurable productivity gains under Vision 2030.

Ethics Approval and Informed Consent

Not applicable. The study uses publicly available aggregate secondary data and includes no human participants, identifiable records, interventions, or private information.

ISSN: 2392- 5418 **e-ISSN:** 2520- 7423 **DOI:**10.59799/SQZH5356

Author Contributions

Conceptualization, methodology, formal analysis, data curation, visualization, writing - original draft preparation, and writing - review and editing: the author. The author approved the submitted version and accepts responsibility for the integrity of the work.

Funding

No specific external funding was reported for this study.

Data Availability Statement

Data on macroeconomic series are publicly available (see the source section in Material and Methods for details) from the World Bank, World Development Indicators. The constructed index, transformed analysis dataset and code can be made available from the corresponding author on reasonable request, subject to the terms of the underlying public sources.

Conflicts of Interest/Competing Interests

The author declares no conflicts of interest or competing interests.

Acknowledgments

None.

ISSN: 2392- 5418 e-ISSN: 2520- 7423 DOI:10.59799/SQZH5356

References

1. Acemoglu D. The simple macroeconomics of AI. Econ Policy. 2025;40(121):13-58. doi:10.1093/epolic/eiae042.

2. Filippucci F, Gal P, Schief M. The impact of artificial intelligence on productivity, distribution and growth: key mechanisms, initial evidence and policy challenges. OECD Artificial Intelligence Papers. 2024;(15). doi:10.1787/8d900037-en.

3. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science. 2023;381(6654):187-92. doi:10.1126/science.adh2586.

4. Brynjolfsson E, Li D, Raymond LR. Generative AI at work. Q J Econ. 2025;140(2):889-942. doi:10.1093/qje/qjae044.

5. Babina T, Fedyk A, He A, Hodson J. Artificial intelligence, firm growth, and product innovation. J Financ Econ. 2024;151:103745. doi:10.1016/j.jfineco.2023.103745.

6. Korinek A. Generative AI for economic research: use cases and implications for economists. J Econ Lit. 2023;61(4):1281-317. doi:10.1257/jel.20231736.

7. Brynjolfsson E, Rock D, Syverson C. The productivity J-curve: how intangibles complement general purpose technologies. Am Econ J Macroecon. 2021;13(1):333-72. doi:10.1257/mac.20180386.

8. Filippucci F, Gal P, Schief M. Miracle or myth? Assessing the macroeconomic productivity gains from artificial intelligence. OECD Artificial Intelligence Papers. 2024;(29). doi:10.1787/b524a072-en.

9. Cazzaniga M, Jaumotte F, Li L, Melina G, Panton AJ, Pizzinelli C, et al. Gen-AI: artificial intelligence and the future of work. IMF Staff Discussion Notes. 2024;2024(1):1-41. doi:10.5089/9798400262548.006.

10. Cerutti EM, Garcia Pascual AI, Kido Y, Li L, Melina G, Mendes Tavares M, et al. The global impact of AI: mind the gap. IMF Working Papers. 2025;2025(76):1-33. doi:10.5089/9798229008570.001.

11. Filippucci F, Gal P, Laengle K, Schief M. Macroeconomic productivity gains from artificial intelligence in G7 economies. OECD Artificial Intelligence Papers. 2025;(41). doi:10.1787/a5319ab5-en.

12. Filippucci F, Gal P, Schief M. Aggregate productivity gains from artificial intelligence: a sectoral perspective. AEA Pap Proc. 2026;116:31-5. doi:10.1257/pandp.20261035.

ISSN: 2392- 5418 **e-ISSN:** 2520- 7423 **DOI:**10.59799/SQZH5356

13. McElheran K, Li JF, Brynjolfsson E, Kroff Z, Dinlersoz E, Foster L, et al. AI adoption in America: who, what, and where. J Econ Manag Strategy. 2024;33(2):375-415. doi:10.1111/jems.12576.

14. Calvino F, Fontanelli L. A portrait of AI adopters across countries: firm characteristics, assets' complementarities and productivity. OECD Science, Technology and Industry Working Papers. 2023;(2). doi:10.1787/0fb79bb9-en.

15. Calvino F, Costa H, Haerle D. Digital technology diffusion in the age of AI: cross-country evidence from microdata. OECD Science, Technology and Industry Working Papers. 2026;(1). doi:10.1787/ebc2debe-en.

16. World Bank. World Development Report 2026: The promise of artificial intelligence. Washington (DC): World Bank; 2026. doi:10.1596/978-1-4648-2331-2.

17. Saudi Vision 2030. Vision 2030 annual report 2025 [Internet]. Riyadh: Saudi Vision 2030; 2026 [cited 2026 Aug 15]. Available from: https://www.vision2030.gov.sa/en/annual-reports

18. Saudi Data and AI Authority. National strategy for data and artificial intelligence [Internet]. Riyadh: SDAIA; 2020 [cited 2026 Aug 15]. Available from: https://sdaia.gov.sa/en/SDAIA/SdaiaStrategies/Documents/NSDAI.pdf

19. Saudi Data and AI Authority. State of AI in Saudi Arabia [Internet]. Riyadh: SDAIA; 2024 [cited 2026 Aug 15]. Available from: https://sdaia.gov.sa/en/MediaCenter/KnowledgeCenter/ResearchLibrary/StateofAIinSaudiArabia.pdf

20. Digital Government Authority. Digital transformation measurement 2025 [Internet]. Riyadh: DGA; 2026 [cited 2026 Aug 15]. Available from: https://dga.gov.sa/en/indicators/qiyas/2025

21. Koll D, Medici A, Mendes Tavares M, Saito M. Structural reforms in Saudi Arabia since 2016. IMF Working Paper WP/26/14. Washington (DC): International Monetary Fund; 2026. doi:10.5089/9798229032728.001.

22. Ahmed EM, Elfaki KE, Alamin YMM, Abdel-Aziz MS, Elmoustafa AMG. Digitalization and climate change spillover effects on Saudi digital economy sustainable economic growth. Fudan J Humanit Soc Sci. 2025;18:687-723. doi:10.1007/s40647-025-00443-6.

23. Barone A, Graupp A, Melhem S. Estimating the economic impacts of AI in Saudi Arabia [Internet]. Washington (DC): World Bank; 2025 Mar 4 [cited 2026 Aug 15]. Available from: https://blogs.worldbank.org/en/arabvoices/estimating-the-economic-impacts-of-ai-in-saudi-arabia

ISSN: 2392- 5418 **e-ISSN:** 2520- 7423 **DOI:**10.59799/SQZH5356

24. World Bank. GDP per person employed (constant 2021 PPP $): metadata glossary, indicator SL.GDP.PCAP.EM.KD [Internet]. Washington (DC): World Bank; 2026 [cited 2026 Aug 15]. Available from: https://databank.worldbank.org/metadataglossary/world-development-indicators/series/SL.GDP.PCAP.EM.KD

25. World Bank. World Development Indicators [Internet]. Washington (DC): World Bank; 2026 [cited 2026 Aug 15]. Available from: https://databank.worldbank.org/source/world-development-indicators

26. General Authority for Statistics. Methodology and quality report for ICT access and usage statistics for establishments 2024 [Internet]. Riyadh: GASTAT; 2026 [cited 2026 Aug 15]. Available from: https://www.stats.gov.sa/en/w/methodology-and-quality-report-for-the-statistics-of-the-survey-of-access-and-use-of-ict-by-establishments-2024

27. General Authority for Statistics. Establishments ICT access and usage statistics 2025 [Internet]. Riyadh: GASTAT; 2026 [cited 2026 Aug 15]. Available from: https://www.stats.gov.sa/documents/d/guest/establishments-ict-access-and-usage-statistics-2025-en-pdf

28. International Monetary Fund. AI Preparedness Index (AIPI): methodology note [Internet]. Washington (DC): IMF; 2024 [cited 2026 Aug 15]. Available from: https://www.imf.org/external/datamapper/AIPINote.pdf

29. World Intellectual Property Organization. Intellectual property statistics: Saudi Arabia country profile [Internet]. Geneva: WIPO; 2026 [cited 2026 Aug 15]. Available from: https://www.wipo.int/edocs/statistics-country-profile/en/sa.pdf

30. Pesaran MH, Shin Y, Smith RJ. Bounds testing approaches to the analysis of level relationships. J Appl Econom. 2001;16(3):289-326. doi:10.1002/jae.616.

31. Sam CY, McNown R, Goh SK. An augmented autoregressive distributed lag bounds test for cointegration. Econ Model. 2019;80:130-41. doi:10.1016/j.econmod.2018.11.001.

32. Bertelli S, Vacca G, Zoia M. Bootstrap cointegration tests in ARDL models. Econ Model. 2022;116:105987. doi:10.1016/j.econmod.2022.105987.

33. Kripfganz S, Schneider DC. ardl: estimating autoregressive distributed lag and equilibrium correction models. Stata J. 2023;23(4):983-1019. doi:10.1177/1536867X231212434.

34. Bi R, Fareed F, Lee JD, Rehman S, Rollinson YG, Yuan T. Digital transformation in the Gulf Cooperation Council economies. IMF Departmental Papers. 2025;2025(3). doi:10.5089/9798400291449.087.

ISSN: 2392- 5418 **e-ISSN:** 2520- 7423 **DOI:**10.59799/SQZH5356

35. World Bank. Digital Progress and Trends Report 2025: Strengthening AI foundations [Internet]. Washington (DC): World Bank; 2025 [cited 2026 Aug 15]. Available from: https://www.worldbank.org/en/publication/dptr2025-ai-foundations