Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Separating AI from Capital: Identifying the Productivity Contribution of Artificial Intelligence across Countries and Sectors

Syed M. Abdullah

s.abullah@lums.edu.pk

Lahore University of Management Sciences (LUMS)

Muhammad Elyoun Seja

KGS

Research Article

Keywords: Artificial intelligence, Total factor productivity, Growth accounting, Productivity J-curve, Shift-share instruments, Penn World Table

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11599318/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Separating AI from Capital: Identifying the Productivity Contribution of Artificial Intelligence across Countries and Sectors

Syed M. Abdullah[1*] and Muhammad Elyoun Seja[2]

[1*]Department of Economics, Lahore University of Management Sciences (LUMS), Lahore, Pakistan.
[2]Department of Economics, KGS, Karachi, Pakistan.

*Corresponding author(s). E-mail(s): s.abullah@lums.edu.pk;
Contributing authors: elyounseja@gmail.com;

Abstract

Does enterprise adoption of artificial intelligence (AI) raise total factor productivity (TFP), or does measured output growth mainly reflect capital deepening and labour growth? We separate these channels in two stages. First, we compute Törnqvist growth-accounting residuals using time-varying labour shares from the Penn World Table (PWT) 10.01, with PWT capital stocks extended to 2024 by perpetual inventory using World Bank investment data. Second, we relate TFP growth to the share of enterprises using AI, taken from Organisation for Economic Co-operation and Development (OECD) business surveys, in a country-year panel with country and year fixed effects. To address endogenous adoption, we use shift-share instruments that interact 2010 fixed-broadband and internet penetration with leave-one-out cross-country AI diffusion, estimated by two-stage least squares (2SLS) and limited-information maximum likelihood (LIML), with weak-instrument-robust Anderson–Rubin inference. Because adoption data cover only 36 countries in recent years (102 country-year observations), statistical power is limited. We find no detectable contemporaneous effect of AI adoption on TFP growth: the fixed-effects estimate is 0.005 percentage points per percentage point of adoption ($\boldsymbol{p = 0.93}$), and the 2SLS and LIML estimates are 0.153 ($\boldsymbol{p = 0.42}$) and 0.299 ($\boldsymbol{p = 0.48}$), with a weak first stage. Point estimates at one- and two-year lags are larger (0.054 and 0.096) but imprecise. A sectoral exposure design with country-by-year and sector-by-year fixed effects, an OECD-only subsample, and a leave-one-country-out jackknife give similar results. The evidence is consistent with, but cannot confirm, a delayed productivity J-curve.

Keywords: Artificial intelligence, Total factor productivity, Growth accounting, Productivity J-curve, Shift-share instruments, Penn World Table

JEL Classification: O33 , O47 , C23 , C26 , E23 , D24

1 Introduction

The rapid diffusion of artificial intelligence (AI), machine learning, and algorithmic automation across enterprise workflows has re-ignited intense debate regarding the macroeconomic trajectory of modern productivity. Proponents argue that generative and analytical AI represents the quintessential general purpose technology (GPT) of the twenty-first century, poised to break decades of secular stagnation by automating cognitive routines, accelerating scientific discovery, and shifting the aggregate production frontier outward (Babina et al. 2024; Brynjolfsson et al. 2017). Conversely, sceptics emphasize that despite pervasive adoption headlines, aggregate total factor productivity (TFP) growth across advanced and emerging economies remains historically subdued, echoing Robert Solow's famous 1987 aphorism that computers could be seen everywhere except in the productivity statistics (Acemoglu 2025; Gordon 2016).

Resolving this dispute empirically at the macroeconomic level presents severe econometric and conceptual difficulties. An observed expansion in gross output coinciding with the installation of enterprise AI systems cannot legitimately be attributed to an AI-driven productivity shock. Output growth may mechanically stem from rapid capital deepening (expenditures on specialized graphics processing units, server racks, cloud computing architecture, and data storage), parallel expansions in skilled labour inputs, or unobserved demand and macroeconomic cycles. Furthermore, national income and product accounts (NIPA) record hardware and software investments as tangible and intangible gross fixed capital formation, while organizational redesign, bespoke algorithmic fine-tuning, workflow adjustments, and internal data curation are rarely capitalized, obscuring the boundary between input accumulation and technological efficiency (Corrado et al. 2021).

This measurement challenge is compounded by two fundamental endogeneity problems. First, in estimating the underlying production elasticities of capital and labour, firms adjust variable inputs in response to productivity shocks observed by managers but unobserved by econometricians—the classical transmission bias identified by Marschak and Andrews (1944). Second, at the secondary stage of estimating the returns to technological adoption, AI adoption itself is highly endogenous: high-performing sectors and frontier economies with superior institutional capacity, higher human capital, and stronger initial digital infrastructure adopt AI faster, generating severe confounding and reverse causality. Naive Ordinary Least Squares (OLS) regressions of productivity on AI metrics inevitably yield biased estimates.

This paper develops an empirical framework that separates the productivity contribution of artificial intelligence from capital and labour deepening. It differs from existing cross-country productivity studies in five respects:

1. **Aggregate Country-Year Panel Architecture with True Capital Stocks:** We estimate primary macroeconomic production and productivity equations at the aggregate country-year level where Penn World Table (PWT 10.01) capital stocks (`rnna`) are directly measured, avoiding artificial sectoral value-added share allocation distortions. To extend PWT `rnna` past 2019, we apply the annual real growth rate of World Bank constant-price gross fixed capital formation (GFCF) to the 2019 PWT real investment series ($I^{\text{PWT}}_{2019} = csh_i_{2019} \times Y^{\text{PWT}}_{2019}$) and accumulate capital via perpetual inventory ($K_t = (1-\delta)K_{t-1} + I_t$).
2. **Cost-Share Törnqvist Residuals and One-Step Joint Estimation:** Rather than computing TFP residuals from econometric production-function regressions, which are imprecise in short macro panels, we compute standard Solow–Jorgenson Törnqvist residuals using time-varying labor shares ($1 - \alpha_{ct} = labsh_{ct}$) from PWT 10.01. We complement this two-step residual approach with a one-step joint structural production function where AI_{ct} enters directly alongside $\ln K$ and $\ln L$.
3. **Country-Year Estimation and a Sectoral Exposure Test:** We estimate Stage 2 regressions at the aggregate country-year level ($N = 102$ across 36 countries) and test micro-mechanisms via a Rajan–Zingales interaction model that interacts national AI adoption with baseline sectoral AI task exposure ($AI_{cst} = AI_{ct} \times \text{Exposure}_s$) while including country-by-year and sector-by-year fixed effects.
4. **Overidentified Instrumental Variables and Weak-Instrument Robustness:** We construct an overidentified system of two predetermined shift-share instruments (Z_1 from 2010 fixed broadband penetration and Z_2 from 2010 internet penetration, each interacted with leave-one-out global AI diffusion). This allows us to estimate Limited Information Maximum Likelihood (LIML) alongside 2SLS, report Hansen J overidentification tests ($df = 1$), and compute Anderson–Rubin weak-instrument-robust confidence tests.
5. **Nested Samples and Transparent Inference:** All subsamples, including the OECD subsample ($N = 92$), are subsets of the main estimation sample ($N = 102$), and inference on the instrumental-variable estimates does not rely on a strong first stage.

Previewing the results, the Solow–Jorgenson factor shares average $\bar{\alpha} = 0.489$ for capital and $\bar{\beta} = 0.511$ for labour. When Törnqvist TFP residuals are regressed on enterprise AI adoption, we find no detectable contemporaneous effect ($\hat{\theta}_{\text{OLS}} = +0.005$, $p = 0.926$; overidentified IV-2SLS $\hat{\theta}_{\text{IV}} = +0.153$, $p = 0.415$; LIML $\hat{\theta}_{\text{LIML}} = +0.299$, $p = 0.483$). The first stage is weak (joint $F = 5.06$), and the Hansen J test does not reject the overidentifying restriction at the 5 percent level ($p = 0.071$). Point estimates at one- and two-year lags are larger ($\hat{\theta}_1 = +0.054$, $\hat{\theta}_2 = +0.096$) but statistically insignificant. An OECD-only subsample and a leave-one-country-out jackknife give similar results.

These estimates should be read in light of the data constraint that motivates much of the design. Enterprise AI adoption is measured for only 36 countries in our estimation sample, mostly in 2020–2024, so the Stage 2 sample contains 102 country-year observations. With country and year fixed effects, identification relies on limited within-country variation, and the confidence intervals are wide enough to be consistent

with both zero and moderate positive effects. We therefore present the results as an early, transparent benchmark rather than a definitive estimate.

The complete code pipeline is available at https://github.com/Syed-Main/ai-productivity-replication/tree/main/code.

The remainder of this article is organized as follows. Section 2 reviews the related literature. Section 3 outlines the theoretical framework. Section 4 details the empirical strategy and instrumental variable identification. Section 5 describes data sources and capital construction. Section 6 presents the empirical results. Section 7 reports the robustness battery. Section 8 discusses policy implications, and Section 9 concludes.

2 Related Literature

Our research intersects four active strands of the economic literature: growth accounting and general purpose technologies, microeconomic evidence on artificial intelligence, macroeconomic growth projections, and the econometrics of panel production functions.

2.1 General Purpose Technologies and Growth Accounting

The foundational architecture for measuring productivity originates from Solow (1957) and was operationalized into comprehensive multi-factor accounting by Jorgenson and Griliches (1967). In this neoclassical growth accounting paradigm, output growth is decomposed into input growth weighted by marginal factor shares, with the unexplained residual interpreted as costless shifts in the aggregate production frontier (total factor productivity).

When applied to major technological transitions, however, standard growth accounting often encounters the paradox identified by Bresnahan and Trajtenberg (1995): general purpose technologies (such as the steam engine, electrification, and digital computing) generate profound economic transformations that fail to appear in concurrent productivity statistics. As Stiroh (2002) and Jorgenson et al. (2008) demonstrated in the context of the 1990s Information and Communication Technology (ICT) revolution, measured productivity accelerations materialized only after sustained capital deepening in computing hardware was paired with sweeping organizational and complementary structural adjustments. Gordon (2016) cautions, however, that digital innovations may exhibit diminishing marginal economic returns compared to the transformative Great Inventions of the second industrial revolution.

2.2 Microeconomic Evidence on Artificial Intelligence

In contrast to macroeconomic aggregates, recent microeconomic and firm-level research documents tangible benefits from AI adoption. Using administrative and job-posting data, Babina et al. (2024) show that AI-investing firms experience faster growth in sales, employment, and market valuation, driven primarily by product innovation rather than cost displacement. Similarly, Brynjolfsson et al. (2025) study the staggered introduction of a generative AI assistant among 5,172 customer-support agents and find an average productivity gain of about 15 percent in issues resolved

per hour, with the largest gains for novice and lower-skilled workers. Alderucci et al. (2020) find that AI-related invention is associated with growth in firm revenue and labour productivity.

Nevertheless, microeconomic productivity premiums do not automatically aggregate into macroeconomic expansions. As Acemoglu and Restrepo (2018, 2020) emphasize in their task-based framework, if technological adoption is characterized by "so-so automation"—substituting capital for labour in routine tasks with only marginal cost reductions without creating new high-productivity tasks—aggregate economic growth remains tepid while labour share declines.

2.3 Macroeconomic Assessments and the Productivity J-Curve

At the aggregate level, Acemoglu (2025) develops a task-based macroeconomic evaluation of AI and estimates that only around 5 percent of tasks are likely to be profitably affected over a ten-year horizon, implying a cumulative TFP gain of no more than about 0.66 percent over ten years (roughly 0.07 percent per year). In complementary work, Filippucci et al. (2024) evaluate OECD cross-country exposure and find substantial heterogeneity based on industrial structure and technological readiness.

The primary explanation for delayed macroeconomic impacts is the *Productivity J-Curve* formalized by Brynjolfsson et al. (2021). They demonstrate that when a new GPT emerges, firms must redirect substantial capital and labour toward creating intangible assets—reorganizing management structures, retraining personnel, accumulating proprietary datasets, and rewiring business workflows. Because national accounts count this diverted effort as input expenditures while omitting the resulting unmeasured intangible capital from output, measured productivity growth initially falls or stagnates. Only in subsequent periods, when intangible capital begins delivering productive services without further adjustment costs, does measured TFP accelerate.

2.4 Identification in Panel Production Functions

Estimating production functions is plagued by endogeneity. Since Marschak and Andrews (1944), econometricians have recognized that profit-maximizing firms adjust input demands in response to anticipated productivity shocks, creating positive correlation between regressors and the error term. While structural proxy variable methods (e.g., Ackerberg et al. 2015; Levinsohn and Petrin 2003; Olley and Pakes 1996; Wooldridge 2009) address this in micro-level firm datasets with intermediate materials data, macro panels spanning heterogeneous countries lack consistent materials data.

Consequently, cross-country productivity research relies on dynamic panel GMM estimators. Arellano and Bond (1991) and Blundell and Bond (1998) developed first-differenced and system GMM estimators that purge unobserved time-invariant individual effects and exploit lagged levels and differences as internal instruments. Roodman (2009) outlines best practices for instrument proliferation and collapsed weighting matrices. For the secondary regression linking technology to TFP, Goldsmith-Pinkham et al. (2020) provide an identification analysis of shift-share (Bartik) instruments, formalizing the orthogonality requirements of initial shares and exogenous aggregate shifters.

3 Theoretical Framework

To provide microfoundations for our empirical design, we integrate the task-based framework of Acemoglu and Restrepo (2018) into an aggregate Solow–Jorgenson growth accounting decomposition.

3.1 Task-Based Production and Technology

Let aggregate output in country c, sector s, and period t be produced via a continuum of tasks $i \in [N-1, N]$:

$$Y_{cst} = \Omega_{cst} \left[\int_{N-1}^{N} y_{cst}(i)^{\frac{\sigma-1}{\sigma}} \, di \right]^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where $\sigma \in (0, \infty)$ is the elasticity of substitution between tasks and Ω_{cst} represents Hicks-neutral baseline efficiency. Each task i can be produced using capital $k_{cst}(i)$ or labour $l_{cst}(i)$ according to the production technology:

$$y_{cst}(i) = \begin{cases} \gamma_K(i) k_{cst}(i) + \gamma_L(i) l_{cst}(i) & \text{if } i \leq I_{cst}, \\ \gamma_L(i) l_{cst}(i) & \text{if } i > I_{cst}, \end{cases} \tag{2}$$

where $I_{cst} \in [N-1, N]$ denotes the threshold of automated tasks. Tasks $i \leq I_{cst}$ are technologically automated and can be executed by algorithmic or physical capital, whereas tasks $i > I_{cst}$ strictly require human labour.

Adoption of artificial intelligence operates through two distinct channels:

1. **Automation effect ($\Delta I_{cst} > 0$):** Expanding the threshold of tasks performed by capital.
2. **Task creation and deepening ($\Delta N_{cst} > 0$, $\Delta \gamma_K > 0$):** Introducing novel complex tasks where labour has a comparative advantage, or deepening capital efficiency in already-automated tasks.

Assuming competitive factor markets with rental rate of capital R_{cst} and wage W_{cst}, cost minimization implies that capital is allocated to tasks where the effective cost of capital is strictly lower than labour: $R_{cst}/\gamma_K(i) < W_{cst}/\gamma_L(i)$. Under constant returns across tasks and Cobb–Douglas aggregation ($\sigma \to 1$), aggregate sectoral output simplifies to the structural function:

$$Y_{cst} = A_{cst} K_{cst}^{\alpha} L_{cst}^{\beta}, \tag{3}$$

where α and β reflect the aggregate output elasticities of capital and labour, and A_{cst} denotes true total factor productivity.

3.2 Intangible Investment and the Productivity J-Curve

Following Brynjolfsson et al. (2021), deploying AI requires unmeasured intangible capital accumulation, denoted J_{cst}. In any period t, a fraction of measured labour and

capital is diverted away from producing final output toward building organizational and digital infrastructure:

$$Y_{cst}^{\text{measured}} = Y_{cst} - \psi(I_{cst}, \Delta AI_{cst}), \quad \psi > 0, \quad \frac{\partial \psi}{\partial \Delta AI} > 0, \tag{4}$$

where $\psi(\cdot)$ represents organizational installation and retraining frictions. The stock of intangible capital accumulates according to $J_{cst} = (1 - \delta_J)J_{cs,t-1} + \lambda AI_{cs,t-1}$. Over time, this unmeasured capital enhances true productivity:

$$\ln A_{cst} = \ln \bar{A}_{cst} + \phi J_{cst}. \tag{5}$$

Taking first differences yields the central testable prediction for measured TFP growth:

$$\Delta \ln A_{cst}^{\text{measured}} = \delta_0 + \theta_0 \Delta AI_{cst} + \sum_{k=1}^{K} \theta_k AI_{cs,t-k} + \boldsymbol{\gamma}' \mathbf{X}_{cst} + \varepsilon_{cst}, \tag{6}$$

where $\theta_0 \leq 0$ captures contemporaneous adjustment frictions and $\theta_k > 0$ for $k \geq 1$ captures the delayed realization of productivity gains.

This theoretical formulation yields three core hypotheses:

- **Hypothesis 1 (J-Curve Trajectory):** Contemporaneous AI adoption has an insignificant or negative association with measured TFP ($\theta_0 \leq 0$), followed by positive lagged returns ($\sum_{k=1}^{K} \theta_k > 0$).
- **Hypothesis 2 (Digital Complementarity):** The productivity elasticity of AI is higher in economies with mature pre-existing digital infrastructure.
- **Hypothesis 3 (Sectoral Divergence):** Due to varying cognitive routine densities, industrial manufacturing and tech-intensive services realize productivity gains faster than primary sectors.

4 Empirical Strategy

Our empirical identification proceeds in two sequential econometric stages, accompanied by a direct one-step joint production function benchmark.

4.1 Stage 1: Production Functions, Factor Shares, and Growth Accounting

4.1.1 Solow–Jorgenson Törnqvist index (primary benchmark)

Rather than deriving residual productivity from econometric production function regressions that are vulnerable to transmission bias and imprecise elasticity estimates, our primary growth accounting benchmark implements the nonparametric Solow–Jorgenson Törnqvist index (Jorgenson and Griliches 1967). Under competitive factor markets and cost-minimizing behavior, the output elasticity of labor equals the cost share of labor, and constant returns to scale holds by construction ($\alpha_{ct} + \beta_{ct} = 1$).

We extract time-varying national labor shares $labsh_{ct}$ directly from the Penn World Table (PWT 10.01), defining:

$$\beta_{ct} = labsh_{ct}, \quad \alpha_{ct} = 1 - labsh_{ct}. \tag{7}$$

The two-period average factor shares are defined as $\bar{\alpha}_{ct} = \frac{1}{2}(\alpha_{ct} + \alpha_{c,t-1})$ and $\bar{\beta}_{ct} = 1 - \bar{\alpha}_{ct}$. Annual output growth is decomposed into the contributions of capital accumulation ($C_{K,ct}$), labor growth ($C_{L,ct}$), and residual TFP growth ($\Delta \ln A_{ct}$):

$$C_{K,ct} = \bar{\alpha}_{ct} \cdot \Delta \ln K_{ct}, \tag{8}$$

$$C_{L,ct} = \bar{\beta}_{ct} \cdot \Delta \ln L_{ct}, \tag{9}$$

$$\Delta \ln A_{ct} = \Delta \ln Y_{ct} - C_{K,ct} - C_{L,ct}. \tag{10}$$

This avoids the returns-to-scale distortions of unrestricted regression residuals, at the cost of assuming competitive factor markets and constant returns to scale.

4.1.2 One-step joint production function

As a direct structural alternative that avoids two-step generated-regressor issues entirely, we estimate a unified one-step production function where enterprise AI adoption enters the output equation directly alongside capital and labor:

$$\ln Y_{ct} = \mu_c + \tau_t + \alpha \ln K_{ct} + \beta \ln L_{ct} + \theta AI_{ct} + \boldsymbol{\gamma}' \mathbf{X}_{ct} + \varepsilon_{ct}, \tag{11}$$

where μ_c is country fixed effects, τ_t is year fixed effects, and θ is the change in log output associated with a one-percentage-point increase in enterprise AI adoption, conditional on capital and labor inputs.

4.1.3 Econometric benchmarks (OLS-FE, IV-FD, System GMM)

For comparative validation against the growth accounting literature, we also estimate standard Cobb–Douglas specifications:

$$\ln Y_{cst} = \mu_{cs} + \tau_t + \alpha \ln K_{cst} + \beta \ln L_{cst} + \omega_{cst} + \nu_{cst}. \tag{12}$$

We estimate equation (12) via:

1. **OLS Fixed Effects:** Within-entity demeaning in unrestricted and per-worker CRS-restricted form ($\ln(Y/L) = \mu_{cs} + \tau_t + \alpha \ln(K/L) + u$).
2. **First-Differenced IV (Anderson–Hsiao):** Differencing to remove μ_{cs} and instrumenting ($\Delta \ln K, \Delta \ln L$) with second-lagged levels ($\ln K_{t-2}, \ln L_{t-2}$).
3. **System GMM (Blundell–Bond):** Stacking differenced and levels equations with collapsed instrument sets ($GMM_MAX_LAGS = 4$).

4.2 Stage 2: Primary Macro Aggregate and Sectoral Mechanism Models

4.2.1 Primary country-year estimation

Our primary Stage 2 specification estimates the effect of enterprise AI adoption on Törnqvist TFP growth at the aggregate country-year level ($N = 102$, 36 countries):

$$\Delta \ln A_{ct} = \delta + \theta AI_{ct} + \boldsymbol{\gamma}' \mathbf{X}_{ct} + \mu_c + \tau_t + \varepsilon_{ct}, \tag{13}$$

where $\mathbf{X}_{ct}$ includes log GDP per capita, gross tertiary education enrolment, research and development expenditure (% of GDP), and trade openness. TFP growth $\Delta \ln A_{ct}$ is measured in percentage points (100× log difference) and AI_{ct} is the percentage of enterprises using AI, so θ is the change in annual TFP growth, in percentage points, associated with a one-percentage-point increase in adoption.

4.2.2 Overidentified instrumental variables

To isolate exogenous variation in enterprise AI adoption and eliminate reverse causality, we construct an overidentified system of two predetermined shift-share instruments:

$$Z_{1,ct} = \text{Broadband}_{c,2010} \times \overline{AI}_{-c,t}, \tag{14}$$

$$Z_{2,ct} = \text{Internet}_{c,2010} \times \overline{AI}_{-c,t}, \tag{15}$$

where $\text{Broadband}_{c,2010}$ and $\text{Internet}_{c,2010}$ measure fixed broadband and internet user penetration in the predetermined baseline year 2010 (prior to the commercial deep learning boom), and $\overline{AI}_{-c,t} = \frac{1}{M_t - 1} \sum_{j \neq c} AI_{jt}$ is the leave-one-out mean adoption rate across the M_t countries observed in year t. Identification rests on the exogeneity of the 2010 penetration shares (Goldsmith-Pinkham et al. 2020); because both shares are strongly correlated with each other (correlation 0.90 across the 36 estimation countries) and with economic development, we treat exclusion as an assumption and discuss it in Section 8.3.

Because the system is overidentified ($L = 2$ instruments for $K = 1$ endogenous variable), we estimate the model using both Two-Stage Least Squares (2SLS) and Limited Information Maximum Likelihood (LIML). Unlike 2SLS, LIML is approximately median-unbiased and less biased when instruments are weak (Stock et al. 2002). Overidentification provides two critical diagnostic tests:

1. **Hansen J Test of Overidentifying Restrictions:** We compute the Hansen J-statistic ($\sim \chi^2(1)$ under H_0), testing whether the two instruments yield mutually consistent estimates. This is a necessary but not sufficient condition for validity: the test cannot detect a violation of exclusion shared by both instruments.
2. **Anderson–Rubin (AR) Weak-Instrument Test:** To ensure robust inference regardless of first-stage strength, we report the Anderson–Rubin test statistic, which maintains exact nominal size and valid confidence sets even under arbitrarily weak instruments (Anderson and Rubin 1949; Stock and Yogo 2005).

Because within R^2 has no natural interpretation for IV and LIML regressions, we do not report it for those columns.

4.2.3 Sectoral exposure model (Rajan–Zingales interaction)

To test microeconomic task exposure mechanisms while controlling for unobserved macroeconomic confounders, we estimate a Rajan–Zingales interaction specification on the country-sector-year panel:

$$\Delta \ln A_{cst} = \theta \left(AI_{ct} \times \text{Exposure}_s \right) + \mu_{ct} + \lambda_{st} + \varepsilon_{cst}, \tag{16}$$

where Exposure_s is a stylised AI task-exposure weight for sector s ($\text{Exposure}_{\text{Services}} = 1.00$, $\text{Exposure}_{\text{Industry}} = 0.60$, $\text{Exposure}_{\text{Agriculture}} = 0.15$). These weights are set by the authors to reflect the ordering of sectors in occupational AI-exposure measures (Felten et al. 2021); they are not estimated, and the results are conditional on them. Specification (16) includes country-by-year fixed effects (μ_{ct}), which absorb country-level macroeconomic, institutional, and policy shocks, and sector-by-year fixed effects (λ_{st}), which absorb global sector-specific trends.

4.2.4 Horizon-specific lag specifications

Because enterprise AI adoption is observed mainly in 2020, 2021, 2023, and 2024, including several lags in a single regression would leave very few observations. To preserve sample coverage while evaluating the Productivity J-Curve hypothesis, we estimate separate horizon-specific panel regressions (analogous to local projections) for each lag horizon $k \in \{0, 1, 2\}$:

$$\Delta \ln A_{ct} = \delta_k + \theta_k AI_{c,t-k} + \boldsymbol{\gamma}_k' \mathbf{X}_{ct} + \mu_c + \tau_t + \varepsilon_{ct}^{(k)}, \quad k \in \{0, 1, 2\}, \tag{17}$$

comparing the contemporaneous association (θ_0) with the lagged associations (θ_1, θ_2). Because each horizon is estimated on a different sample, we do not sum the coefficients into a cumulative effect.

5 Data Architecture and Capital Construction

The analysis combines three international sources: the Penn World Table (Feenstra et al. 2023), the World Bank World Development Indicators (World Bank 2026), and the OECD survey of information and communication technology (ICT) use in businesses (OECD 2026). The aggregate country-year panel covers 180 countries over 2000–2024 (4,236 observations), and the sectoral panel covers 166 countries and three sectors (Agriculture, Industry, and Services) over the same period (11,606 country-sector-year observations; both panels are unbalanced). Enterprise AI adoption is available for 40 countries (174 country-year observations between 2017 and 2025, concentrated in 2020–2024). After requiring non-missing controls, the Stage 2 estimation sample contains 102 country-years in 36 countries, of which 90 observations are high-income economies. Country-average adoption in this sample ranges from about 5 to

29 percent of enterprises (median about 12 percent). Table 1 summarizes the core variables, series codes, and transformation rules.

Variable	Concept	Source and identifier	Construction and definition
Y_{cst}	Real Sectoral Output	World Bank WDI	Real value added by sector (constant 2015 US$): `NV.AGR.TOTL.KD`, `NV.IND.TOTL.KD`, `NV.SRV.TOTL.KD`.
L_{cst}	Sectoral Labour	World Bank WDI	Total employment multiplied by sectoral employment share (%): $\text{Emp}_{ct} \times (\text{Share}_{cst}/100)$, where Emp = LabourForce$\times(1-$Unemp$/100)$.
K_{ct}, K_{cst}	Capital Stock	PWT 10.01 + WB GFCF	National capital stock at constant 2017 prices (`rnna`) extended post-2019 by splicing 2019 PWT real investment ($csh_i \times rgdpna$) with annual real WB GFCF growth (`NE.GDI.FTOT.KD`); allocated by value-added share in the sectoral mechanism panel.
$labsh_{ct}$	Labour Cost Share	PWT 10.01	Share of labour compensation in GDP at current national prices (`labsh`), used to construct Törnqvist factor weights $\beta_{ct} = labsh_{ct}$ and $\alpha_{ct} = 1 - labsh_{ct}$.
AI_{ct}	AI Adoption Rate	OECD SDMX (`DSD_ICT_B`)	Percentage of enterprises employing artificial intelligence technologies (`G14_B`, all activities and enterprise sizes).
$\text{Broadband}_{c,2010}$	Shift-Share Base (Z_1, Z_2)	World Bank WDI	Predetermined 2010 fixed broadband subscriptions per 100 people (`IT.NET.BBND.P2`) and 2010 internet user penetration (`IT.NET.USER.ZS`), interacted with leave-one-out global AI diffusion.
$\ln \text{GDPpc}_{ct}$	Economic Development	World Bank WDI	Natural log of GDP per capita, constant 2015 US$ (`NY.GDP.PCAP.KD`).
Tertiary_{ct}	Human Capital	World Bank WDI	Gross tertiary education enrolment ratio (% gross, `SE.TER.ENRR`).
RD_{ct}	Innovation Intensity	World Bank WDI	Research and development expenditure as a percentage of GDP (`GB.XPD.RSDV.GD.ZS`).
Trade_{ct}	Trade Openness	World Bank WDI	Trade volume (% of GDP): (Exports + Imports)/GDP (`NE.TRD.GNFS.ZS`).
Internet_{ct}	Digital Density	World Bank WDI	Percentage of individuals using the Internet (`IT.NET.USER.ZS`).

Table 1 Data sources, indicators, and variable construction.

5.1 Capital Stock Construction and Perpetual Inventory Extension

A fundamental challenge in cross-country productivity research is constructing consistent capital stocks. Prior sectoral studies frequently allocated aggregate national capital to sectors in proportion to employment shares ($K_{cst} = K_{ct} \times (L_{cst}/L_{ct})$), which mathematically forces capital-to-labour ratios to be identical across all sectors within a country-year ($K_{cst}/L_{cst} = K_{ct}/L_{ct}$). While allocating capital by value-added shares relaxes this restriction, any sectoral allocation rule remains an imputation.

Our empirical strategy therefore proceeds along two complementary paths:

1. **Aggregate Country-Year Benchmark (Primary):** We estimate our primary macroeconomic production functions and Stage 2 TFP equations at the aggregate country-year level, where the capital stock is directly measured by the Penn World Table (PWT 10.01, `rnna`) without requiring arbitrary sectoral allocation assumptions. PWT `rnna` reflects asset-specific perpetual inventory calculations across ten distinct physical and intangible asset categories at constant 2017 national prices (Feenstra et al. 2015).

   Because official PWT 10.01 coverage concludes in 2019, extending capital stocks through 2024 requires a perpetual inventory extension. Rather than scaling World Bank investment by the arbitrary ratio K_{2019}/I_{2019} (which conflates capital-output ratios with investment price deflators), we establish baseline PWT real investment in 2019 directly from national accounts shares:

$$I_{c,2019}^{\text{PWT}} = csh_i_{c,2019} \times Y_{c,2019}^{\text{PWT}}, \tag{18}$$

   where csh_i is the investment share of GDP at current PPPs and Y^{PWT} is real GDP (`rgdpna`). For subsequent years $t > 2019$, real investment evolves strictly according to the annual real growth rate of World Bank constant-price gross fixed capital formation (I_{ct}^{WB}):

$$I_{c,t}^{\text{PWT}} = I_{c,t-1}^{\text{PWT}} \times \left(\frac{I_{c,t}^{\text{WB}}}{I_{c,t-1}^{\text{WB}}} \right), \quad \forall t \geq 2020. \tag{19}$$

   Capital stocks then accumulate via the perpetual inventory equation:

$$K_{c,t} = (1 - \delta_c) K_{c,t-1} + I_{c,t}^{\text{PWT}}, \tag{20}$$

   where δ_c is country c's empirical depreciation rate from PWT 10.01 (averaging 5.0 percent). Real GDP and employment are extended using analogous World Bank real growth rates.
2. **Sectoral Mechanism Panel (Secondary):** For our secondary Rajan–Zingales mechanism analysis, national capital is distributed across sectors according to gross value-added shares ($K_{cst} = K_{ct} \times [Y_{cst}/\sum_j Y_{cjt}]$). Crucially, under this imputation $\Delta \ln K_{cst} = \Delta \ln Y_{cst} + \Delta \ln(K_{ct}/Y_{ct})$, so the aggregate capital-output ratio component $\bar{\alpha}_{ct} \Delta \ln(K_{ct}/Y_{ct})$ is nonparametrically absorbed by the country-by-year fixed

effects (μ_{ct}) in specification (16), so that within-country-year identification comes from sectoral labour-productivity differentials weighted by the national labour share ($\bar{\beta}_{ct}\Delta \ln(Y_{cst}/L_{cst})$). The sectoral growth decomposition in Section 6 inherits this imputation, so sectoral capital contributions partly reflect the allocation rule rather than observed sectoral investment.

6 Empirical Results

6.1 Stage 1: Production Function and Factor-Share Estimates

Table 2 presents the empirical factor shares and structural production function estimates across the Törnqvist benchmark, one-step joint model, OLS-FE, IV in first differences, and System GMM.

Specification / Parameter	**Estimate**	**Std. Error**
α (Törnqvist Cost-Share)	0.489	0.002
β (Törnqvist Cost-Share)	0.511	0.002
$\alpha + \beta$ (Törnqvist Cost-Share)	1.000	—
α (One-Step Joint)	0.074	0.147
β (One-Step Joint)	0.441	0.084
$\alpha + \beta$ (One-Step Joint)	0.515	—
θ_{AI} (One-Step Joint)	4.6×10^{-5}	2.0×10^{-4}
α (OLS unrestricted)	0.487	0.042
β (OLS unrestricted)	0.235	0.087
$\alpha + \beta$ (OLS unrestricted)	0.722	—
α (OLS CRS imposed)	0.509	0.042
$\beta = 1 - \alpha$ (OLS CRS imposed)	0.491	0.042
α (IV-FD, Anderson–Hsiao)	1.411	0.297
β (IV-FD, Anderson–Hsiao)	0.121	0.645
$\alpha + \beta$ (IV-FD)	1.532	—
α (System GMM, Blundell–Bond)	0.025	0.067
β (System GMM, Blundell–Bond)	0.120	0.080
$\alpha + \beta$ (System GMM)	0.145	—

Table 2 Stage 1: Production-function and factor-share estimates across models. Standard errors for the Törnqvist shares are standard errors of the sample mean

The empirical labor share in PWT 10.01 averages $\bar{\beta} = 0.511$ ($SE = 0.002$), implying an average capital elasticity of $\bar{\alpha} = 0.489$, with constant returns to scale ($\alpha + \beta = 1.000$) imposed by construction. This is our benchmark for growth accounting. In the one-step joint estimation, where enterprise AI adoption enters the aggregate production function directly alongside capital and labor (second panel of Table 2), the coefficient on AI is $\hat{\theta}_{\mathrm{AI}} = 4.6 \times 10^{-5}$ (SE $= 2.0 \times 10^{-4}$, $p = 0.815$, $N = 102$), so

we detect no direct output shift beyond conventional inputs. The one-step input elasticities are imprecise and imply strongly decreasing returns ($\hat{\alpha} + \hat{\beta} = 0.515$), which reflects the small sample; we therefore rely on the cost-share residuals as the primary productivity measure.

Among the econometric benchmarks, OLS-FE yields an unrestricted capital elasticity of $\hat{\alpha} = 0.487$ ($SE = 0.042$) and labour elasticity of $\hat{\beta} = 0.235$ ($SE = 0.087$), with CRS rejected (Wald stat $= 10.57, p = 0.0011$). Under per-worker CRS, the capital elasticity is $\hat{\alpha} = 0.509$ ($SE = 0.042$). The first-differenced IV estimator yields an imprecise sum ($\hat{\alpha}+\hat{\beta} = 1.532$), reflecting weak instruments when inputs are differenced on short panels. In the dynamic System GMM model, the autoregressive parameter on lagged output is $\hat{\rho} = 0.889$ ($p < 0.001$), while short-run capital and labour elasticities are 0.025 and 0.120, respectively (Hansen test $\chi^2(9) = 17.46, p = 0.042$; AR(1) $z = -2.88, p = 0.004$; AR(2) $z = 0.09, p = 0.926$). The implausibly low System GMM elasticities (sum 0.145) and the rejection of the Hansen test at 5 percent are consistent with weak internal instruments in short macro panels; we report these estimates for comparison only.

6.2 Growth Accounting Decomposition

Table 3 and Figure 1 display the mean annual growth accounting decomposition using Törnqvist factor shares over the 2000–2024 period.

Sector	**Output (%)**	**Capital (pp)**	**Labour (pp)**	**TFP (pp)**
Agriculture	1.98	1.06	-0.19	1.11
Industry	3.07	1.58	0.97	0.52
Services	3.96	2.04	1.37	0.55

Table 3 Mean annual sectoral growth decomposition using Törnqvist cost shares (percentage points, 2000–2024).

The Törnqvist cost-share decomposition reveals pronounced structural divergence across sectors alongside positive baseline efficiency gains. Services exhibited the fastest annual real output growth (3.96 percentage points), driven by capital accumulation (2.04 pp contribution), labour force expansion (1.37 pp contribution), and residual TFP growth (0.55 pp). Industry grew at 3.07 pp annually, supported by capital deepening (1.58 pp), labour absorption (0.97 pp), and TFP growth (0.52 pp). In Agriculture, output grew at 1.98 pp, characterized by structural labour outflow (-0.19 pp contribution) more than offset by capital deepening (1.06 pp) and productivity gains (1.11 pp). Grounding factor weights in observed PWT cost shares avoids the scale distortions of unrestricted regression residuals. Because sectoral capital is imputed from value-added shares (Section 5), the sectoral split between capital and TFP contributions should be read as indicative.

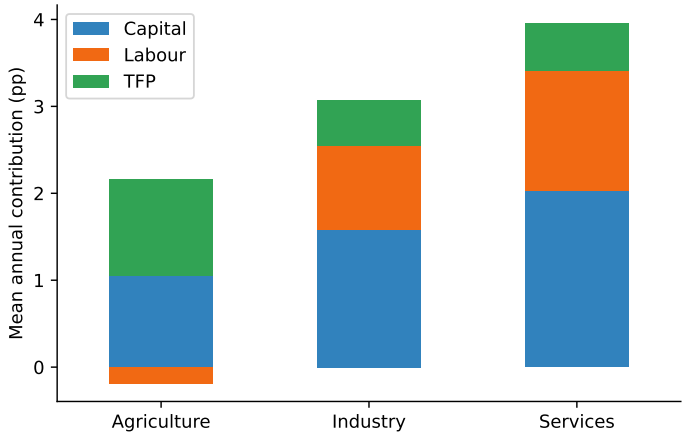


Fig. 1 Mean annual contribution of capital, labour, and TFP growth by sector (percentage points, 2000–2024)

Table 4 Stage 2: AI adoption and TFP growth (aggregate and sectoral specifications)

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS-FE	OLS-FE (lag 1)	OLS-FE (lag 2)	2SLS	LIML	Sectoral R–Z (CY+SY FE)
AI adoption (θ)	0.0048	0.0541	0.0965	0.1531	0.2989	0.0255
	(0.0516)	(0.0940)	(0.0945)	(0.1878)	(0.4262)	(0.0368)
Controls	Yes	Yes	Yes	Yes	Yes	No
Country FE	Yes	Yes	Yes	Yes	Yes	—
Year FE	Yes	Yes	Yes	Yes	Yes	—
Country × Year FE	No	No	No	No	No	Yes
Sector × Year FE	No	No	No	No	No	Yes
Observations	102	68	65	102	102	399
Countries	36	35	34	36	36	37
Within R^2	0.419	0.628	0.285	—	—	0.001
First-stage F	—	—	—	5.06	5.06	—
Hansen J (p-value)	—	—	—	0.071	0.086	—
Anderson–Rubin (p-value)	—	—	—	0.136	0.136	—

Notes: Standard errors clustered at the entity level in parentheses. No coefficient is statistically significant at the 10 percent level. The dependent variable is annual Törnqvist TFP growth in percentage points; AI adoption is the percentage of enterprises using AI. Columns (1)–(5) estimate country-year panel equations; column (6) estimates the Rajan–Zingales (R–Z) sectoral exposure equation with country-by-year (CY) and sector-by-year (SY) fixed effects. Within R^2 is not reported for the IV and LIML columns. Instruments Z_1 and Z_2 are 2010 fixed-broadband and internet penetration interacted with leave-one-out mean AI adoption.

6.3 Stage 2: AI Adoption and TFP Growth

Table 4 reports the Stage 2 econometric estimations examining the relationship between enterprise AI adoption and residual TFP growth across aggregate country-year specifications and the sectoral Rajan–Zingales mechanism model.

In the baseline contemporaneous OLS-FE model (Column 1), the estimated AI coefficient is $\hat{\theta} = +0.0048$ ($SE = 0.0516, p = 0.926, N = 102$), indicating that enterprise AI adoption is not associated with an immediate expansion in annual TFP growth.

In the instrumental-variable specifications (Z_1 from 2010 broadband penetration and Z_2 from 2010 internet penetration), the joint first-stage F-statistic is 5.06, below the rule-of-thumb threshold of 10 (Staiger and Stock 1997). Because the two 2010 penetration shares are highly correlated, the second instrument adds little independent variation. With a weak first stage, 2SLS can be biased toward OLS and conventional confidence intervals can be misleading, so we also report LIML and weak-instrument-robust Anderson–Rubin inference. The IV-2SLS estimate (Column 4) is $\hat{\theta}_{\text{IV}} = +0.1531$ ($SE = 0.1878, p = 0.415$), while the LIML estimator (Column 5) yields $\hat{\theta}_{\text{LIML}} = +0.2989$ ($SE = 0.4262, p = 0.483$).

The diagnostics are as follows:

1. **Hansen J Test:** The overidentification test yields $J = 3.268$ ($p = 0.071$) for 2SLS and $J = 2.956$ ($p = 0.086$) for LIML. The overidentifying restriction is not rejected at the 5 percent level but is rejected at the 10 percent level. Because the test has low power when instruments are weak and cannot detect a violation of exclusion shared by both instruments, it does not establish instrument validity.
2. **Anderson–Rubin Test:** The Anderson–Rubin test of $H_0 : \theta = 0$ yields $F = 2.042$ ($p = 0.136$). This test has correct size regardless of instrument strength, so the absence of a significant effect is not an artefact of 2SLS size distortion. Its power is limited in this sample, however, so moderate positive effects cannot be ruled out.
3. **Rajan–Zingales Sectoral Exposure:** In Column (6), interacting national AI adoption with sectoral task exposure, with country-by-year and sector-by-year fixed effects, yields $\hat{\theta} = +0.0255$ ($SE = 0.0368$, $p = 0.489$, $N = 399$). Within countries and years, sectors more exposed to AI do not show detectably faster TFP growth as national adoption rises.

6.4 Lagged Associations and the Productivity J-Curve

Table 5 and Figure 2 report the horizon-specific estimates used to evaluate Hypothesis 1 across horizons $k \in \{0, 1, 2\}$.

Lag Horizon (k)	$\hat{\theta}_k$	**Std. Error**	p**-value**	N
$k = 0$ (Current)	0.0048	0.0516	0.9259	102
$k = 1$ (1 Year)	0.0541	0.0940	0.5704	68
$k = 2$ (2 Years)	0.0965	0.0945	0.3179	65

Table 5 Horizon-specific coefficients of enterprise AI adoption on TFP growth (separate OLS-FE regressions for each horizon $k \in \{0, 1, 2\}$; samples differ across horizons)

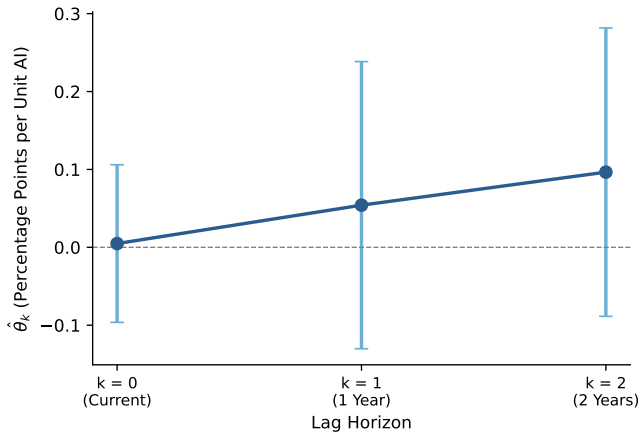


Fig. 2 Horizon-specific coefficients ($\hat{\theta}_k$) of AI adoption on TFP growth from separate regressions, with 95 percent confidence intervals

The point estimates rise from $\hat{\theta}_0 = +0.0048$ ($SE = 0.0516$, $N = 102$) at $k = 0$ to $\hat{\theta}_1 = +0.0541$ ($SE = 0.0940$, $N = 68$) at one lag and $\hat{\theta}_2 = +0.0965$ ($SE = 0.0945$, $N = 65$) at two lags (Table 4, Columns 2 and 3, and Table 5). This ordering is what the Productivity J-Curve of Brynjolfsson et al. (2021) would predict if gains arrive only after complementary intangible investment. However, none of the coefficients is statistically significant, the 95 percent confidence intervals overlap almost entirely (Figure 2), and the samples differ across horizons, so the pattern cannot be distinguished from a flat profile. The contemporaneous estimate is also positive rather than negative, so Hypothesis 1's prediction of initial frictions ($\theta_0 \leq 0$) is not supported either. Longer adoption series will be needed to test the J-curve directly.

6.5 Heterogeneity Decomposition: Income Groups and Sectors

Table 6 and Figure 3 present the heterogeneity analysis decomposing the AI–TFP relationship across economic development tiers and productive sectors.

Subsample Group	$\hat{\theta}$	**Std. Error**	p**-value**	N
Income Tiers:				
High-Income Economies	-0.0096	0.0545	0.8609	90
Productive Sectors:				
Services	-0.0227	0.0549	0.6802	99
Industry	0.0108	0.0664	0.8716	99
Agriculture	0.0303	0.1731	0.8616	99

Table 6 Heterogeneity of the AI adoption coefficient across income groups and sectors

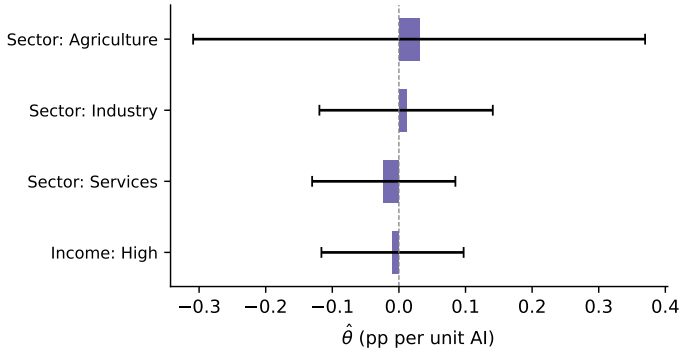


Fig. 3 Estimated AI coefficient ($\hat{\theta}$) by income group and sector, with 95 percent confidence intervals

The heterogeneity estimates are informative mainly about what the data cannot show:

1. **Income groups:** High-income economies account for 90 of the 102 estimation observations (31 countries). Their estimate is close to zero ($\hat{\theta} = -0.0096$, $SE = 0.0545$). Because only 12 observations come from upper-middle-income economies and none from lower-income economies, we do not estimate separate effects for those groups, and the data do not allow a test of Hypothesis 2.
2. **Sectors:** Across the three sectors ($N = 99$ each, 35 countries), none of the estimates is distinguishable from zero or from the others (Services -0.0227, Industry $+0.0108$, Agriculture $+0.0303$). The sign pattern should not be interpreted as evidence for Hypothesis 3.

7 Robustness Battery

To verify the empirical validity of our baseline findings, we conduct four robustness checks, summarized in Table 7.

Diagnostic / Robustness Test	**Value**	**Interpretation / Empirical Check**
Translog Production $\alpha(K)$	0.3977	Output elasticity of capital at sample means.
Translog Production $\beta(L)$	0.3173	Output elasticity of labour at sample means.
Translog Scale Returns $(\alpha + \beta)$	0.7150	Sum of elasticities at sample means.
OECD-Only Subsample $\hat{\theta}$	-0.0033	$SE = 0.0504$; $N = 92$, 32 countries.
Placebo Test (p-value)	0.8326	Randomly generated pseudo-adoption, pre-2014 sample; $\hat{\theta}_{\text{fake}} = -0.0197$.
Jackknife Mean $\hat{\theta}$	0.0058	Mean of 36 leave-one-country-out estimates.
Jackknife Standard Error	0.0615	Range of leave-one-out estimates: $[-0.0378, 0.0297]$.

Table 7 Robustness checks

7.1 Translog production function

To examine whether the Stage 1 results depend on the Cobb–Douglas assumption of a unitary elasticity of substitution ($\sigma = 1$), we estimate a generalized transcendental logarithmic (translog) specification:

$$\ln Y_{ct} = \mu_c + \tau_t + \alpha_1 \ln K + \alpha_2 \ln L + \alpha_3 (\ln K)^2 + \alpha_4 (\ln L)^2 + \alpha_5 (\ln K \times \ln L) + u_{ct}. \quad (21)$$

Evaluated at sample means, the elasticities are $\alpha_K = 0.3977$ and $\beta_L = 0.3173$ (sum $= 0.7150$). The second-order terms matter: the coefficient on $(\ln K)^2$ is -0.0286 ($p = 0.023$) and the interaction coefficient α_5 is 0.0514 ($p = 0.078$), so Cobb–Douglas is a simplification. The translog elasticities also imply decreasing returns, whereas the Törnqvist benchmark imposes constant returns; Stage 2 results based on the Törnqvist residuals should be read with this caveat.

7.2 OECD subsample

Restricting the baseline OLS-FE model to OECD members leaves 92 observations in 32 countries, a subset of the full estimation sample ($N = 102$, 36 countries). The estimate is $\hat{\theta} = -0.0033$ ($SE = 0.0504$), again statistically indistinguishable from zero.

7.3 Placebo check

We assign randomly generated pseudo-adoption values to the pre-2014 sample and re-estimate the Stage 2 model. The placebo coefficient is $\hat{\theta}_{\text{fake}} = -0.0197$ ($p = 0.8326$). This confirms that the estimation pipeline does not mechanically generate spurious coefficients. Because the pseudo-treatment is random by construction, however, the check is not a test for pre-existing trends correlated with actual adoption.

7.4 Leave-one-country-out jackknife

To check whether individual countries drive the results, we drop each of the 36 countries in turn and re-estimate the baseline model. The jackknife mean is $\hat{\theta}_{\text{JK}} = 0.0058$ with jackknife standard error 0.0615, and the leave-one-out estimates range from -0.0378 to 0.0297, so no single country drives the estimate.

8 Discussion and Policy Implications

8.1 Reconciling Macro Insignificance with Micro Optimism

Our finding of no detectable aggregate contemporaneous TFP effect ($\hat{\theta}_{\text{OLS}} = +0.005$, IV-2SLS $\hat{\theta}_{\text{IV}} = +0.153$, IV-LIML $\hat{\theta}_{\text{LIML}} = +0.299$) contrasts with firm-level studies that document sizeable gains for early AI adopters (Babina et al. 2024; Brynjolfsson et al. 2025). Three mechanisms could reconcile the two, although our data cannot distinguish between them:

1. **Business Stealing versus Frontier Expansion:** If early adopters mainly capture market share from non-adopting rivals, firm-level gains overstate aggregate

gains. More generally, in task-based aggregation the macroeconomic effect depends on the share of tasks affected and the cost savings per task, which may both be small at present (Acemoglu 2025).

2. **Limited Aggregate Task Exposure:** Even with high enterprise adoption percentages, AI systems currently automate only a narrow subset of routine cognitive tasks. In our estimation sample, country-average enterprise adoption ranges from about 5 to 29 percent, and even within adopting firms AI may affect only a small share of tasks.
3. **Unmeasured Intangible Capital Divergence:** In accordance with Corrado et al. (2021) and Brynjolfsson et al. (2021), substantial productive resources are currently absorbed by organizational restructuring, data cleaning, and workflow design. Because these activities represent intermediate costs under standard accounting rules rather than final output, measured productivity is temporarily depressed.

8.2 Implications

With the caveats above, the findings have three implications for policymakers and national statistical agencies:

- **Realistic Expectations for Near-Term Gains:** Estimates near zero in advanced economies ($\hat{\theta} = -0.0033$ in the OECD subsample; $\hat{\theta} = -0.0096$ in high-income economies) suggest that near-term aggregate productivity gains from adoption alone are likely to be small. Complementary investments in skills and organization may matter, although our data cannot test this directly.
- **Modernizing National Accounts for Intangibles:** Statistical agencies should accelerate efforts to capitalize intangible data assets and organizational investments. Without accurate measurement of intangible capital accumulation, productivity statistics will continue to misdiagnose the early stages of technological revolutions as economic stagnation.
- **Task Creation over Displacement:** If gains are delayed, as the imprecise lag estimates hint, policies that support complementary organizational investment and the creation of new tasks (Acemoglu and Restrepo 2018) may matter more for aggregate productivity than adoption rates alone.

8.3 Limitations

Several limitations qualify these results. First, the Stage 2 sample is small (102 country-years, mostly high-income economies, 2020–2024), which limits power and external validity. Second, the instruments are weak, and their exclusion restriction is an assumption: 2010 broadband and internet penetration could affect later TFP growth through general ICT channels other than AI adoption. Third, the Törnqvist residuals impose constant returns and competitive factor markets, whereas the translog estimates point to decreasing returns. Fourth, the sectoral exposure weights are stylised rather than estimated, and sectoral capital is imputed from value-added shares. Finally, enterprise adoption rates measure whether firms use AI, not how intensively. Longer and broader adoption series would address the first and last of these directly.

9 Conclusion

This study provides a cross-country evaluation of the productivity contribution of artificial intelligence that separates AI adoption from capital and labour deepening. Combining Penn World Table 10.01 capital stocks, OECD enterprise AI adoption data, and World Bank indicators, we employ a two-stage design featuring Solow–Jorgenson Törnqvist factor share residuals ($\bar{\alpha} = 0.489, \bar{\beta} = 0.511$), an overidentified instrumental variable system (Z_1 from 2010 broadband penetration and Z_2 from 2010 internet penetration), LIML estimation with Hansen J and Anderson–Rubin diagnostics, and a Rajan–Zingales sectoral task-exposure mechanism model.

Our central conclusion is that, in the 36 countries with enterprise adoption data, AI adoption has not yet produced a detectable acceleration in aggregate total factor productivity ($\hat{\theta}_{\text{OLS}} = +0.005$, IV-2SLS $\hat{\theta}_{\text{IV}} = +0.153$, IV-LIML $\hat{\theta}_{\text{LIML}} = +0.299$, all statistically indistinguishable from zero). Point estimates at one and two lags are larger ($\hat{\theta}_1 = +0.054$, $\hat{\theta}_2 = +0.096$) but imprecise. These results are consistent with an early adoption phase of a general purpose technology, but given the short and small sample they cannot confirm a productivity J-curve. As adoption series lengthen and cover more countries, the framework and replication code provided here can be re-estimated to test whether a delayed aggregate productivity gain emerges.

Supplementary Information. The replication package, including harmonized datasets, factor decomposition routines, regression pipelines, and diagnostic scripts, is available at https://github.com/Syed-Main/ai-productivity-replication.

Acknowledgements. The authors thank seminar participants and colleagues for helpful discussions and constructive feedback.

Statements and Declarations

- **Funding**: No external funding was received for conducting this study.
- **Conflict of Interest / Competing Interests**: The authors declare that they have no competing interests, financial or non-financial, directly or indirectly related to the work submitted.
- **Ethics Approval and Consent to Participate**: Not applicable. This study analyzes publicly available aggregate country- and sector-level data and involves no human participants or animals.
- **Consent for Publication**: Not applicable.
- **Data Availability**: All data used in this study are publicly available: Penn World Table version 10.01 (Feenstra et al. 2023), https://www.rug.nl/ggdc/productivity/pwt/; World Bank World Development Indicators (World Bank 2026), https://databank.worldbank.org/source/world-development-indicators; and the OECD ICT Access and Usage by Businesses database (OECD 2026), https://data-explorer.oecd.org. The harmonized datasets generated during the study are available in the replication repository listed under Code Availability.
- **Code Availability**: The replication code is available at https://github.com/Syed-Main/ai-productivity-replication (code directory: https://github.com/Syed-Main/ai-productivity-replication/tree/main/code).

- **Author Contribution**: Both authors contributed equally to conceptualization, econometric modeling, data curation, empirical analysis, manuscript writing, and editing.

Appendix A Data Cleaning, Winsorization, and Sample Balance

All monetary variables drawn from the World Bank (value added, gross fixed capital formation, GDP per capita) are denominated in constant 2015 US dollars, while Penn World Table capital stocks (`rnna`) and real GDP (`rgdpna`) are measured at constant 2017 national prices and extended beyond 2019 by splicing 2019 PWT real investment ($I^{\text{PWT}}_{c,2019} = csh_i_{c,2019} \times Y^{\text{PWT}}_{c,2019}$) with annual real World Bank GFCF growth rates via perpetual inventory with country-specific depreciation rates δ_c. To limit the influence of extreme measurement noise and macroeconomic crises, estimated Törnqvist TFP growth residuals ($\Delta \ln A_{ct}$ and $\Delta \ln A_{cst}$) are winsorized at the 1st and 99th percentiles ($WINSOR_PCTILES = (0.01, 0.99)$). Entities with non-positive values for real output, capital, or labour are removed prior to logarithmic transformation.

Appendix B Computational Pipeline and Code Architecture

The empirical pipeline is written in Python and organized in modules (code directory: https://github.com/Syed-Main/ai-productivity-replication/tree/main/code). The pipeline is structured across 11 dedicated modules:

1. `config.py`: Global parameter definitions, API endpoints, time windows, and winsorization thresholds.
2. `utils.py`: Cached REST ingestion utilities, console-to-disk logging tee, and panel demeaning functions.
3. `api_worldbank.py`: Automated batch extraction of WDI sectoral value added, employment, and macroeconomic controls.
4. `api_oecd.py`: OECD SDMX REST parser extracting enterprise AI adoption rates (`DSD_ICT_B`).
5. `api_pwt.py`: Automated download and parsing of Penn World Table 10.01 capital stocks (`rnna`) and human capital.
6. `panel.py`: Harmonization of international series into balanced panel with value-added capital allocation and Bartik instrument construction.
7. `stage1.py`: Structural production function estimation (OLS-FE, IV-FD, and System GMM) and Solow residual decomposition.
8. `stage2.py`: Regression of TFP growth on enterprise AI adoption (OLS-FE, shift-share 2SLS, LIML, and horizon-specific lags).
9. `robustness.py`: Comprehensive battery (translog production function, OECD subsample, placebo pre-period test, and leave-one-out jackknife).
10. `reporting.py`: Automated generation of publication-quality booktabs LaTeX tables, CSV files, and vector PDF figures.

11. `run_analysis.py`: Master pipeline orchestrator supporting live API updates and offline validation.

Appendix C Software Dependencies and Replication Protocol

The econometric analysis is implemented in Python 3.12 utilizing `pandas` (2.2.3), `numpy` (2.1.3), `linearmodels` (7.0), `statsmodels` (0.14.4), `scipy` (1.14.1), `pydynpd` (0.4.0), `requests` (2.32.3), and `openpyxl` (3.1.5). The replication package is available at https://github.com/Syed-Main/ai-productivity-replication. Running `python run_analysis.py` from within the `code/` directory reproduces all empirical tables and figures directly.

References

Acemoglu D (2025) The simple macroeconomics of AI. Economic Policy 40(121):13–58. https://doi.org/10.1093/epolic/eiae042

Acemoglu D, Restrepo P (2018) The race between man and machine: Implications of technology for growth, factor shares, and employment. American Economic Review 108(6):1488–1542

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

Ackerberg DA, Caves K, Frazer G (2015) Identification properties of recent production function estimators. Econometrica 83(6):2411–2451

Alderucci D, Branstetter L, Hovy E, et al (2020) Quantifying the impact of AI on productivity and labor demand: Evidence from U.S. Census microdata. Paper presented at the Allied Social Science Associations Annual Meeting, San Diego, URL https://www.aeaweb.org/conference/2020/preliminary/paper/Tz2HdRna

Anderson TW, Rubin H (1949) Estimation of the parameters of a single equation in a complete system of stochastic equations. Annals of Mathematical Statistics 20(1):46–63

Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58(2):277–297

Babina T, Fedyk A, He A, et al (2024) Artificial intelligence, firm growth, and product innovation. Journal of Financial Economics 151:103745

Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87(1):115–143

Bresnahan TF, Trajtenberg M (1995) General purpose technologies: 'Engines of growth'? Journal of Econometrics 65(1):83–108

Brynjolfsson E, Rock D, Syverson C (2017) Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. NBER Working Paper 24001, National Bureau of Economic Research

Brynjolfsson E, Rock D, Syverson C (2021) The productivity J-curve: How intangibles complement general purpose technologies. American Economic Journal: Macroeconomics 13(1):333–372

Brynjolfsson E, Li D, Raymond L (2025) Generative AI at work. Quarterly Journal of Economics 140(2):889–942. https://doi.org/10.1093/qje/qjae044

Corrado C, Haskel J, Jona-Lasinio C (2021) Artificial intelligence and productivity: An intangible assets approach. Oxford Review of Economic Policy 37(3):435–458

Feenstra RC, Inklaar R, Timmer MP (2015) The next generation of the Penn World Table. American Economic Review 105(10):3150–3182

Feenstra RC, Inklaar R, Timmer MP (2023) Penn World Table version 10.01. Groningen Growth and Development Centre, University of Groningen, URL https://www.rug.nl/ggdc/productivity/pwt/

Felten E, Raj M, Seamans R (2021) Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. Strategic Management Journal 42(12):2195–2217. https://doi.org/10.1002/smj.3286

Filippucci F, Gal P, Jona-Lasinio C, et al (2024) The impact of artificial intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges. OECD Artificial Intelligence Papers, OECD Publishing, Paris

Goldsmith-Pinkham P, Sorkin I, Swift H (2020) Bartik instruments: What, when, why, and how. American Economic Review 110(8):2586–2624

Gordon RJ (2016) The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War. Princeton University Press, Princeton

Jorgenson DW, Griliches Z (1967) The explanation of productivity change. Review of Economic Studies 34(3):249–283

Jorgenson DW, Ho MS, Stiroh KJ (2008) A retrospective look at the U.S. productivity growth resurgence. Journal of Economic Perspectives 22(1):3–24. https://doi.org/10.1257/jep.22.1.3

Levinsohn J, Petrin A (2003) Estimating production functions using inputs to control for unobservables. Review of Economic Studies 70(2):317–341

Marschak J, Andrews WH (1944) Random simultaneous equations and the theory of production. Econometrica 12(3/4):143–205

OECD (2026) ICT access and usage by businesses (dataflow `DSD_ICT_B`). OECD Data Explorer, Paris, URL https://data-explorer.oecd.org

Olley GS, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. Econometrica 64(6):1263–1297

Roodman D (2009) How to do xtabond2: An introduction to difference and system GMM in Stata. Stata Journal 9(1):86–136

Solow RM (1957) Technical change and the aggregate production function. Review of Economics and Statistics 39(3):312–320

Staiger D, Stock JH (1997) Instrumental variables regression with weak instruments. Econometrica 65(3):557–586

Stiroh KJ (2002) Information technology and the U.S. productivity revival: What do the industry data say? American Economic Review 92(5):1559–1576

Stock JH, Yogo M (2005) Testing for weak instruments in linear IV regression. In: Andrews DWK, Stock JH (eds) Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge University Press, Cambridge, p 80–108

Stock JH, Wright JH, Yogo M (2002) A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20(4):518–529

Wooldridge JM (2009) On estimating firm-level production functions using proxy variables to control for unobservables. Economics Letters 104(3):112–114

World Bank (2026) World development indicators. World Bank, Washington, DC, URL https://databank.worldbank.org/source/world-development-indicators