Comparative Advantage in AI-Intensive Industries: Evidence from US Imports

Alessandra Bonfiglioli[1] Rosario Crinò[2] Mattia Filomena[3] Gino Gancia[4]

EEA Congress 2025

[1]Uni Bergamo, QMUL and CEPR
[2]Uni Bergamo, CEPR and CESifo
[3]Uni Bergamo
[4]Uni Milano-Bicocca and CEPR

The Global Race to Harness AI

- AI and Machine Learning can transform manufacturing and service industries
  - applications: predictive maintenance, quality control, customization, supply chain management, inventory optimization, pricing algorithms, customer services...
- research on AI-ML started in the 1950s
  - business applications since late 1990s thanks to computing power and big data
- global race to harness AI is still underway
  - yet AI adoption varies dramatically across sectors and countries
- beg the question:
  - what are the sources of Comparative Advantage (CA) in AI-intensive industries?

This Paper

- following Romalis (2004) and Chor (2010), we test:
  - which country-level factors raise exports to the US in AI-intensive industries
  - focus on exports to the US in 1999-2019
    - ★ variation in AI intensity across 79 industries (4-digit NAICS)
    - ★ variation in sources of CA across (up to) 68 origin countries
- measure AI intensity as share of "AI workers" in employment
- we identify 4 sources of CA
  - endowments: STEM graduates (or ICT capital)
  - digital infrastructure
  - economies of scale
  - (less) regulation on digital trade
- extensive robustness checks, including (historical) IV for STEM graduates

Model

- generalized version of Eaton-Kortum (2002)
  - N countries, F factors, I industries with varieties $j \in [1,0]$
- value of industry-i US imports from country c

$$\ln M_c^i = \underbrace{\theta \sum_{f=1}^{F} \left(\ln \frac{V_{cf}}{V_{c0}} \right) s_f^i}_{H-O} + \underbrace{\theta \ln A_c^i}_{Ricardian} + \theta \ln \tau_c^i + \delta_c + \delta_i$$

  - V_{cf} = factor f endowment of country c
  - s_f^i = factor f intensity of industry i
  - A_c^i = average productivity of country c in industry i
- further assumption on A_c^i in AI-intensive industries:
  - A_c^i increasing in digital infrastructure (I_c) and scale (X_c = total export)

$$A_c^i = \alpha \cdot s_{AI}^i \cdot I_c + \beta \cdot s_{AI}^i \cdot X_c$$

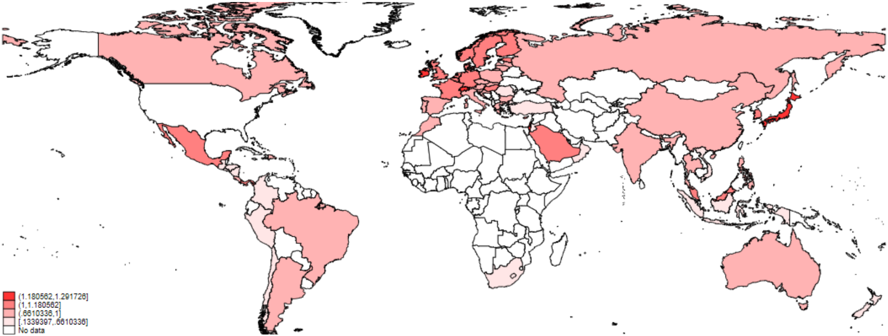
The Data

- US imports
  - manufacturing (Feenstra, Romalis & Schott, 2002)
    - product-level data → 66 industries (NAICS 4-digit)
  - services (BEA) → 13 industries
- sources of comparative advantage
  - digital skills → 45 countries
    - population share of STEM graduates (OECD)
  - digitial infrastructure → 68 countries
    - share of population with Internet access (WDI)
  - economies of scale → 68 countries
    - total export volume (WDI)
  - regulation (later)
    - Digital Services Trade Restrictiveness Index - DSTRI (OECD)

AI Intensity

- employment share of AI-related occupations in each industry in the US
  - from Bonfiglioli et al. (2025)
- industry employment
  - US Census (2000) and ACS (2010, 2020)
- AI-related occupations
  - based on software and skill requirements in US job posting (O*NET)
  - 19 occupations for which:
    - at least two ML or data analysis software are "in demand"
    - require skills in ICT, maths or statistics
- use industry ranking in 2020
  - top: advanced services including public sector
  - bottom: traditional manufacturing

Revealed CA in AI-Intensive Industries

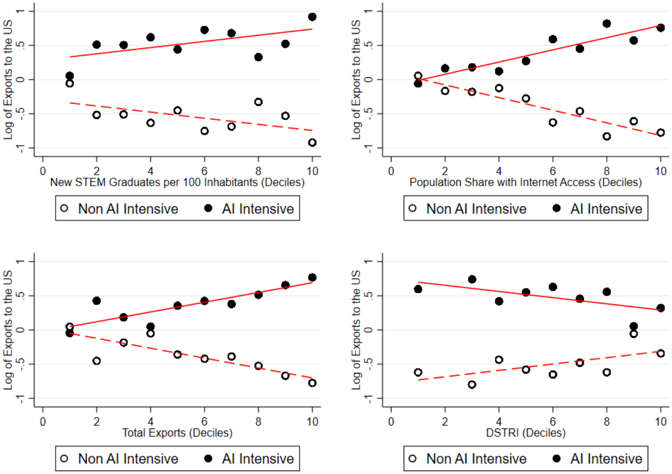


$$RCA_c^{AI} = (M_c^{AI} / M_c^{total}) / (M_{world}^{AI} / M_{world}^{total})$$

- M_c^i = US imports in industry i from country c

Preliminary Evidence

Exports to US and country characteristics by AI intensity



The Empirical Strategy: Specification

- estimation equation:

$$\ln M_{cit} = \alpha_{ct} + \alpha_{it} + \beta \cdot (Z_{ct} \times AIint_i^r) + \boldsymbol{\Gamma}_{cit}\gamma + \varepsilon_{cit},$$

  - M_{cit} = US imports, origin c, industry i, year t
  - Z_{ct} = source of CA in AI
  - $AIint_i^r$ = industry ranking by AI intensity (in 2020)
  - $\boldsymbol{\Gamma}_{cit}$ = controls
  - α_{ct} = country $\times$ year FE
  - α_{it} = industry $\times$ year FE

- coefficient of interest:

$$\beta \text{ (CA if } \beta > 0)$$

- two-way clustering at country-year and industry-year level

Results: Baseline Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$S_{ct} \times AIint_i^r$	0.073***		0.094***			0.057**	0.051**
	(0.014)		(0.022)			(0.022)	(0.022)
$I_{ct} \times AIint_i^r$				0.063***		0.040***	0.027***
				(0.005)		(0.008)	(0.009)
$X_{ct} \times AIint_i^r$					0.008***	0.004***	0.004***
					(0.001)	(0.001)	(0.001)
$G_{ct} \times AIint_i^r$		0.012***	-0.006			-0.004	-0.004
		(0.003)	(0.005)			(0.005)	(0.005)
Controls	No	No	No	No	No	No	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	49,812	49,887	49,742	93,531	93,458	49,535	49,535
Adj. R^2	0.658	0.657	0.658	0.629	0.625	0.662	0.663

- S_{ct}= STEM graduates
- I_{ct}= digital infrastructure
- X_{ct}= total export
- G_{ct}= total graduates

Robustness and Threats to Identification

- robustness:
  - alternative measures of $AIint_i^r$ (Webb, Hansen) or endowments (ICT)
  - excluding extreme obs.
  - excluding AI producing industries
  - alternative specifications (zeros, $\ln AIint_i$, non-parametric)
  - alternative clustering
- threats to identifications:
  - additional interactions (more CA, other controls $\times$ AIint_i^r)
  - additional FEs (country-industry trends, bins-year FEs)

Reverse Causality

- can changes in exports to US $\longrightarrow$ changes in CA?
  - *AIint* measured for US + stable ranking over time
  - but endowments are endogenous
- initial CA:
  - interact $AIint^r_{i,2000}$ with $Z_{c,1999}$ (both predetermined)
  - cross-sectional estimates for period-average of $\ln M_{cit}$
- IV for digital skills endowment:
  - education in STEM result of culture/attitudes toward science
  - instrument = # famous scientists born in c over the period 1780-1880
    - ⋆ from De la Croix and Licando (2015)
    - ⋆ interacted with $AIint^r_{i,2000}$

Reverse Causality: Results

	OLS				2SLS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$S_c \times AIint_i^r$	0.097***			0.061*	0.639***	0.700**	0.820***	1.565**
	(0.028)			(0.031)	(0.121)	(0.345)	(0.167)	(0.690)
$I_c \times AIint_i^r$		0.069***		0.024*		0.064**		0.118**
		(0.015)		(0.013)		(0.029)		(0.052)
$X_c \times AIint_i^r$			0.006***	0.005***		-0.001		-0.009
			(0.002)	(0.001)		(0.004)		(0.007)
$G_c \times AIint_i^r$	-0.023***			-0.008	-0.131***	-0.142*	-0.169***	-0.323**
	(0.008)			(0.008)	(0.027)	(0.073)	(0.036)	(0.144)
First-Stage Regression								
$T_c \times AIint_i^r$					0.009***	0.003***	0.007***	0.002***
					(0.000)	(0.001)	(0.001)	(0.000)
Kleibergen-Paap $F-$Statistic					651.84	26.65	169.31	20.02
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,499	5,070	5,070	3,499	2,263	2,263	2,263	2,263
Adj. R^2	0.689	0.688	0.689	0.692	0.700	0.697	0.675	0.516

Role of Regulation

	Regulation Areas						
	All	All	Infrastructure & Connectivity	Electronic Transactions	E-Payment Systems	Intellectual Property Rights	Other Barriers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$D_{ct} \times AIint_i^r$	-0.040**	-0.049**	-0.044*	0.133	-0.128	-0.492	-0.177*
	(0.016)	(0.023)	(0.024)	(0.112)	(0.165)	(0.318)	(0.098)
$S_{ct} \times AIint_i^r$		0.039	0.037	0.040	0.043*	0.025	0.048**
		(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.024)
$I_{ct} \times AIint_i^r$		0.038**	0.041**	0.048***	0.047***	0.049***	0.042***
		(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)
$X_{ct} \times AIint_i^r$		0.003**	0.003**	0.003**	0.004**	0.003**	0.004**
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$G_{ct} \times AIint_i^r$		-0.003	-0.002	-0.002	-0.003	-0.000	-0.005
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23,076	16,366	16,366	16,366	16,366	16,366	16,366
Adj. R^2	0.631	0.671	0.671	0.670	0.670	0.671	0.670

- additional CA source: $D_{ct} \times AIint_i^r$
    - $D_{ct} \in [0,1]$ = Digital Services Trade Restrictiveness Index (OECD)

Conclusions

- what determines CA in AI-intensive industries?
- we identify 4 sources of CA
  - digital skills (STEM graduates and ICT capital)
  - digital infrastructure (internet access)
  - economies of scale
  - (less restrictive) regulation of digital services
- more questions ahead
  - is gen-AI different?
  - redistributive effects of AI trade towards STEM workers?
  - beyond trade: AI effects on aggregate performance

Related Literature

- sources of CA
  - factor endowment: Romalis (2004), Nunn (2007), Manova (2013), Bonfiglioli et al. (2019)
  - productivity: Eaton and Kortum (2002), Chor (2010), Costinot et al. (2012)
  - new: focus on CA in AI-intensive industries
- effects of new technologies on trade
  - AI and trade in mobile app services: Sun&Trefler (2022, 2023), Brynjolfsson et al. (2019)
  - internet and trade: Blum and Goldfarb (2006), Lendle et al. (2016), Chen andWu (2021), Demir et al. (2025)

Quantification: Baseline Regressions

- cross-country variation in CA sources:
  - S_{75pc} (France) $\rightarrow S_{25pc}$ (Mexico) $\Longrightarrow \ln\left(M_{75}/M_{25}\right) = 0.27$
    - ★ M_{75} = Other General Purpose Machinery Manufacturing
    - ★ M_{25} = Metal Forging and Stamping
  - $I_{75pc} \rightarrow I_{25pc} \Longrightarrow \ln\left(M_{75}/M_{25}\right) = 0.32$
  - $\ln X_{75pc} \rightarrow \ln X_{25pc} \Longrightarrow \ln\left(M_{75}/M_{25}\right) = 0.26$
- average time variation (1999-2019) in CA sources:
  - $\Delta S = 0.12 \rightarrow \ln\left(M_{75}/M_{25}\right) = 0.23$
  - $\Delta I = 0.69 \rightarrow \ln\left(M_{75}/M_{25}\right) = 0.71$
  - $\Delta \ln X = 0.81 \rightarrow \ln\left(M_{75}/M_{25}\right) = 0.12$

Alternative Measures

	AI Intensity				STEM Fields				Digital Skills
	$AIint_i$	STEM Occupations (All)	STEM Occupations (No Life/Medical/Other)	Webb's Occupational AI Exposure	Natutal Sciences	Mathematics & Statistics	Engineering	ICT	ICT Share of Capital Stock
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$S_{ct} \times AIint_i^r$	1.1050*	0.044**	0.047**	0.081***	-0.036	0.205*	0.046*	0.188***	0.254**
	(0.537)	(0.022)	(0.022)	(0.022)	(0.039)	(0.118)	(0.028)	(0.056)	(0.124)
$I_{ct} \times AIint_i^r$	0.687***	0.039***	0.038***	0.016*	0.031***	0.029***	0.027***	0.030***	0.046***
	(0.203)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.011)
$X_{ct} \times AIint_i^r$	0.068**	0.007***	0.007***	0.004***	0.005***	0.004***	0.004***	0.004***	0.004***
	(0.029)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$G_{ct} \times AIint_i^r$	-0.039	-0.008	-0.008	-0.019***	0.004	0.003	0.001	-0.004	
	(0.126)	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)	(0.004)	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	49,535	49,535	49,535	49,535	40,706	46,487	43,946	47,114	39,964
Adj. R^2	0.663	0.667	0.667	0.663	0.655	0.655	0.662	0.656	0.691

Sources of CA: Descriptives

	New STEM Graduates per 100 Inhabitants (S_{ct})			Population Share with Internet Access (I_{ct})			Total Exports (X_{ct})		
	Mean	SD	Δ Mean	Mean	SD	Δ Mean	Mean	SD	Δ Mean
Total	0.237	0.101	0.120	0.491	0.295	0.676	210,049.1	320,351.7	160,966.1
Europe and Central Asia	0.244	0.102	0.138	0.589	0.264	0.703	222,774.2	282,176.5	153,351.0
East Asia and Pacific	0.286	0.082	0.082	0.464	0.304	0.610	374,806.2	488,004.6	357,710.8
America and Caribbean	0.154	0.058	0.058	0.348	0.269	0.617	90,850.1	140,504.6	48,876.3
Africa and Middle East	0.232	0.083	0.083	0.399	0.299	0.812	68,869.8	73,649.8	42,480.8

AI-Related Occupations

Computer and Information Research Scientists	Mathematicians
Computer Network Architects	Network and Computer Systems Administrators
Computer Network Support Specialists	Operations Research Analysts
Computer Occupations, All Other	Software Developers
Computer Programmers	Software Quality Assurance Analysts and Testers
Computer Systems Analysts	Statistical Assistants
Computer User Support Specialists	Statisticians
Data Scientists	Web and Digital Interface Designers
Database Administrators	Web Developers
Database Architects	

Occupations are classified according to the 6-digit level of the 2018 Standard Occupational Classification.

AI Intensity Across Industries

Top Ten Industries	$AIint_i$	Bottom Ten Industries	$AIint_i$
Information and Data Processing Services	0.2633	Dairy Products	0.0089
Computer Equipment Manufacturing	0.2190	Seafood and Other Miscellaneous Foods, nec	0.0088
Communications, Audio and Video Equipment	0.1415	Sawmills and Wood Preservation	0.0087
Other Business Services	0.1363	Miscellaneous Nonmetallic Mineral Products	0.0086
Telecommunication Services	0.1305	Other Woods Products	0.0085
Electronic Components and Products	0.1015	Cement, Concrete, Lime, and Gypsum Products	0.0074
Navigational Electronic and Control Instruments Manufacturing	0.1000	Animal Slaughtering and Processing	0.0070
Financial Services	0.0953	Bakeries	0.0031
Audiovisual Services	0.0915	Construction	0.0030
Aerospace Products Manufacturing	0.0871	Fiber, Yarn, and Thread Mills	0.0028

- top industries: advanced services including public sector
- bottom industries: traditional manufacturing

Imports Trends: AI-intensive Industries

	Above vs. Below Median	Top vs. Bottom Terciles	Top vs. Bottom Quintiles
	(1)	(2)	(3)
AI-Intensive ($\mathbb{I}_i$)	0.235***	0.409***	0.528***
	(0.048)	(0.070)	(0.090)
Country FE	Yes	Yes	Yes
Observations	5,203	3,466	2,061
Adj. R^2	0.157	0.149	0.167

Robustness: Outliers

	AI Intensity		Country Characteristics							Balanced Samples	
	Top & Bottom Five industries (1)	AI Producers (2)	STEM Graduates (3)	Internet Coverage (4)	Total Exports (5)	Real GDP (6)	Real Per Capita GDP (7)	Import Penetration (8)	US FDI (9)	Countries with all Industries (10)	Industries with all Countries (11)
$S_{ct} \times AIint_i^r$	0.052**	0.053**	0.043*	0.056**	0.063**	0.057**	0.057**	0.047**	0.047*	0.064**	0.052*
	(0.026)	(0.024)	(0.025)	(0.025)	(0.026)	(0.024)	(0.022)	(0.023)	(0.026)	(0.026)	(0.028)
$I_{ct} \times AIint_i^r$	0.030***	0.029***	0.021**	0.031***	0.023**	0.021**	0.025***	0.034***	0.023**	0.025**	0.053**
	(0.011)	(0.010)	(0.009)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.020)
$X_{ct} \times AIint_i^r$	0.006***	0.005***	0.004***	0.004***	0.005***	0.004**	0.004***	0.005***	0.005***	0.003*	0.002
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
$G_{ct} \times AIint_i^r$	-0.002	-0.005	0.002	-0.003	-0.007	-0.003	-0.004	-0.011**	-0.001	-0.005	-0.001
	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	44,035	47,319	45,076	42,945	43,250	44,119	44,983	43,984	40,253	20,145	6,840
Adj. R^2	0.666	0.664	0.668	0.654	0.634	0.642	0.671	0.673	0.658	0.696	0.725

Robustness: Sample Selection

	Inverse Hyperbolic Sine Transformation	Heckman Correction	Poisson Pseudo-Maximum Likelihood
	(1)	(2)	(3)
$S_{ct} \times AIint_i^r$	0.045**	0.053**	0.144***
	(0.019)	(0.022)	(0.037)
$I_{ct} \times AIint_i^r$	0.023***	0.028***	-0.016
	(0.008)	(0.009)	(0.021)
$X_{ct} \times AIint_i^r$	0.007***	0.003**	0.006**
	(0.001)	(0.001)	(0.003)
$G_{ct} \times AIint_i^r$	-0.003	-0.005	-0.003
	(0.004)	(0.005)	(0.011)
Controls	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Observations	51,184	49,535	51,184
Adj. R^2	0.696	0.663	0.793

Observable Confounders

	Other Determinants of CA	Additional Controls × $AIint_i^r$					Additional Controls × Industry FE				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$S_{ct} \times AIint_i^r$	0.073***	0.051**	0.057***	0.045**	0.050**	0.050**	0.051**	0.062***	0.045**	0.050**	0.050**
	(0.026)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)	(0.022)	(0.021)	(0.022)	(0.022)	(0.023)
$I_{ct} \times AIint_i^r$	0.027**	0.026***	0.009	0.026***	0.026***	0.023**	0.025***	0.015	0.026***	0.027***	0.023**
	(0.011)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)
$X_{ct} \times AIint_i^r$	0.005***	0.004**	0.003**	0.004***	0.004***	0.005***	0.005***	0.003***	0.004***	0.004***	0.005***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
$G_{ct} \times AIint_i^r$	-0.010	-0.005	-0.005	-0.003	-0.004	-0.002	-0.004	-0.005	-0.004	-0.005	-0.002
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	36,991	49,535	49,535	49,535	49,313	44,236	49,535	49,535	49,535	49,313	44,236
Adj. R^2	0.653	0.663	0.664	0.664	0.664	0.651	0.674	0.683	0.691	0.664	0.659

Threats: Trends and Contemporaneous Shocks

	Underlying Trends		Contemporaneous Shocks		
	(1)	(2)	(3)	(4)	(5)
$S_{ct} \times AIint_i^r$	0.047**	0.049**	0.054**	0.051**	0.051**
	(0.022)	(0.023)	(0.021)	(0.022)	(0.022)
$I_{ct} \times AIint_i^r$	0.027***	0.025***	0.025***	0.028***	0.027***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
$X_{ct} \times AIint_i^r$	0.003**	0.003**	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$G_{ct} \times AIint_i^r$	-0.003	-0.002	-0.003	-0.003	-0.004
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Controls	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
Observations	49,535	48,386	49,535	49,535	49,535
Adj. R^2	0.668	0.659	0.685	0.665	0.663