Artificial Intelligence and the Direction of Firm Innovation

How Predictive AI Shapes Inventive Search

Cher Li[1]
Xin Deng[2], Maria Garcia-Vega[3], Richard Upward[3]

TPI Academic Conference (Sept 2026)

[1]Aston Business School
[2]Alliance Manchester Business School
[3]University of Nottingham

The Puzzle: Exploration vs. Exploitation

View A: Exploration (The Hope)

- **Concept:** AI as a General Purpose Technology (GPT)
- **Logic:** Reduces barriers to entry; lowers cost of scanning combinatorial spaces
- **Examples:** AlphaFold (Biology), Material Science discovery

View B: Exploitation (The Counter-Theory)

- **Concept:** AI as a Prediction Machine
- **Logic:** Works best with abundant training data and well-defined objectives
- **Examples:** Demand forecasting, supply chain coordination, "data exhaust" analysis

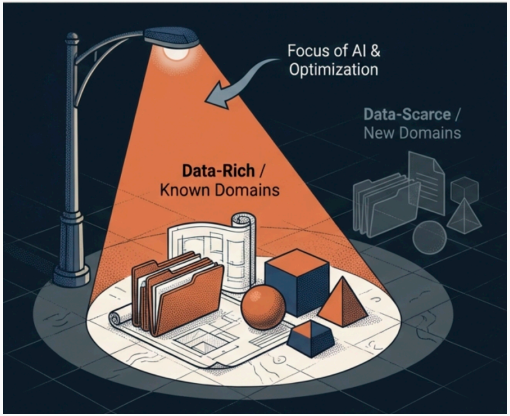
The Mechanism: Data-Visibility Bias

Inventive Search in the Predictive Era

1. **Data Dependence:** Predictive AI requires training data.
2. **The "Local" Advantage:** Firms possess deep "data exhaust" on current processes.
3. **The Streetlight Effect:** AI lowers the cost of local search where the light is strongest.
4. **The Opportunity Cost:** As returns to local optimisation rise, the cost of exploring the dark increases.

"Prediction technologies... raise the opportunity cost of exploration in data-scarce domains."

— Gans (2025)



Data-Visibility Bias in Inventive Search

Data: Matching Skills to Innovation (2017–2021)

Data Sources & Matching

Data Source A: Innovation

Orbis IP (UK Patenting firms, 2010–2021).
Metric: Patents filed with UK IPO, EPO, or USPTO.

Data Source B: AI Input

Adzuna (Near-universe of UK online job vacancies).
Volume: ~1million vacancies per snapshot.

The Match

4,747 unique UK patenting firms matched.
"AI Adopter" = Firms posting vacancies with keywords like "Machine Learning", "NLP", "Neural Networks".

Rising Proportion of AI Vacancies

Defining the Direction of Search

Known Patent: Filed in an IPC class where the firm has patented in the last 5 years.

New Patent: Filed in an IPC class that is completely *new* to the firm's history.

Empirical Strategy: Entropy Balancing + DiD

1. **The Challenge:** Selection Bias. AI adopters are larger, richer and more innovative *ex-ante*.
2. **The Solution:** Entropy Balancing.
   - Reweights control group to match treated group on **mean and variance** of pre-treatment characteristics (Patents, Employment, Assets).
3. **Estimation (PPML):**

$$E[y_{it}] = \exp(\theta(AI_i \times Post_t) + \alpha_i + \delta_t)$$

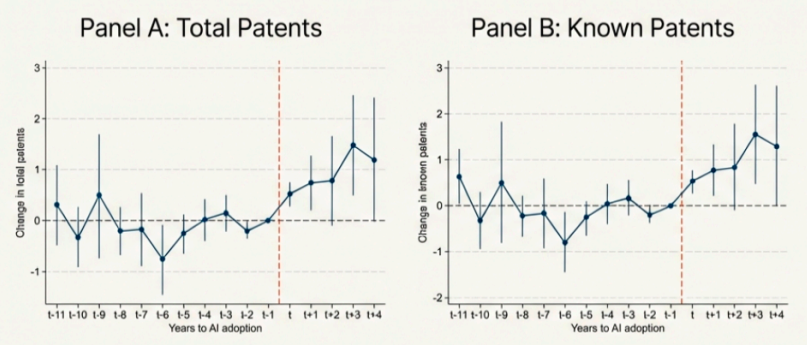
- y_{it}: Count of patents (Total, Known or New).
- α_i, δ_t: Firm and Year fixed effects.

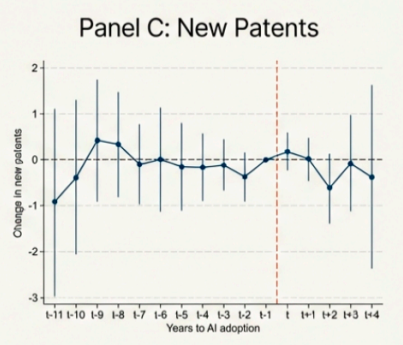
Main Finding: AI Fuels Exploitation

Metric	Effect Size	Significance
Total Patents	+18.4% to +47.1%	**Yes**
Known Domains (Exploitation)	**+18.5% to +47.3%**	**Yes**
New Domains (Exploration)	Statistically Insignificant	No

Conclusion: AI acts as an **Interpolation Technology**, filling in gaps within the known knowledge frontier, rather than an Extrapolation engine.

Event Studies: A Tale of Two Trajectories





Robustness and Measurement Checks

Identification Robustness

- **Alternative Matching**: Results hold using Propensity Score Matching (PSM) and Coarsened Exact Matching (CEM) instead of Entropy Balancing.
- **Falsification Tests**: Pre-treatment trends show no anticipation effects; "placebo" job-posting dates yield null results.
- **Sample Restrictions**: Findings consistent when excluding the "Big Tech" sector or highly skewed patenters.

Measurement Checks

- **AI Skill Dictionary**: Results robust to using alternative keyword lists (e.g., OECD vs. extended dictionary).
- **Innovation Direction**: Tested alternative definitions of "New Domains" (e.g., 3-digit vs. 4-digit IPC classes).
- **Model Sensitivity**: Estimates consistent across Poisson, Negative Binomial and OLS with log-transformed variables.

Mechanism 1: Managerial vs. R&D: Does It Matter *Where* AI Capabilities Are Located

Managerial AI Capabilities (Control)

- **Roles**: Managers & Senior Officials.
- **Logic**: AI improves coordination, monitoring and operational forecasting.
- **Effect**: Reinforces focus on **Known** domains (Optimization).

Scientific AI Capabilities (Augmentation)

- **Roles**: Scientists, Engineers, R&D specialists.
- **Logic**: AI lowers the cost of experimentation and knowledge recombination.
- **Effect**: Broadens search into **New** domains (Discovery).

Managerial AI
(Optimization)

Professional/Scientific AI
(Augmentation)

Mechanism 1: Occupational Locus of AI Skills

Estimation Results: PPML Regressions (Table 11)

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
Panel A: Managers or senior officials with AI skills			
Managers with $AI_i \times$ Post	0.887***	0.949***	−0.664
	(0.309)	(0.347)	(0.623)
N	14092	6988	12419
Panel B: Professionals with AI skills			
Professionals with $AI_i \times$ Post	0.430***	0.427***	0.360*
	(0.104)	(0.106)	(0.218)
N	14068	6988	12400

- **Managers:** Significant only for *Known* patents.
- **Scientists:** Positive and significant for *both* Known and New.

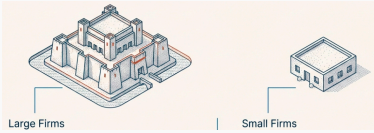
Note: PPML estimates after entropy balancing. Standard errors clustered at firm-level.

Mechanism 2: Firm Size and Complementary Assets

The "Data Moat" Effect

- **Logic:** AI requires real-time data on markets, customers and competitors to be effective.
- **Large Firms:** Possess deeper "data exhaust," established market presence and stronger organizational capacity.
- **Finding:** Large firms translate AI capabilities into both exploration and exploitation more effectively than small firms.

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
$AI_i \times \text{Post}$	0.478***	0.479***	0.420*
	(0.139)	(0.143)	(0.221)
$AI_i \times \text{Post} \times \text{Small}$	−1.970***	−2.006***	−1.576***
	(0.461)	(0.511)	(0.573)
N	14071	6990	12403



The Agglomeration of
New Innovations by Large Firms

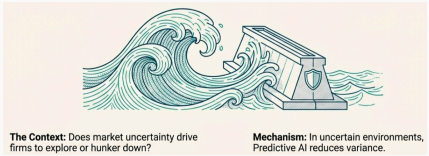
Source: Table 12 - Triple interaction with firm assets.

Mechanism 3: The Role of Market Volatility

Optimisation under Uncertainty

- **Logic:** AI helps analyse consumer behaviour and demand fluctuations in volatile markets.
- **The Tilt:** In uncertain environments, AI provides higher returns to local optimisation (the "safe bet").
- **Finding:** High volatility strengthens the shift toward **Exploitation** in familiar domains.

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
$AI_i \times \text{Post}$	0.340***	0.339***	0.476
	(0.090)	(0.088)	(0.295)
$AI_i \times \text{Post} \times \text{High volatility}$	0.357***	0.371***	−0.195
	(0.121)	(0.114)	(0.347)
N	14071	6990	12403



Predictive AI as a Shield Against Market Volatility

Source: Table 13 - Triple interaction with sector volatility.

Discussion: Predictive Era vs. Generative Future

Predictive Era (2017–2021) *[This Study]*	**Generative Era (Post-2022)** *[Future Outlook]*
Supervised Learning	LLMs / GenAI
Screening, Ranking, optimising	Hypothesising, Drafting
Effect: Narrowing of search ("Data-visibility bias")	**Potential:** Lowering cost of generating *new* hypotheses - but deployment is only starting

Baseline Finding: In the pre-generative regime, AI primarily deepened technological trajectories rather than broadening them.

Conclusion and Policy Implications

Summary

- AI is not a magic wand for discovery. In the predictive era, it deepened existing technological trajectories (**Path Dependence**).
- Firms rationally ignored data-scarce domains to harvest low-hanging fruit in data-rich domains.

Policy Implications

1. **Support Data Generation:** Move the ”streetlight” to novel fields.
2. **Human Capital:** Train scientists to use AI for experimentation/opportunities, not just managers for efficiency.
3. **Competition:** Be wary of ”Data Moats” that entrench incumbents.

Thank You

Cher Li
Aston Business School
c.li@aston.ac.uk