Academic Conference | Technology for Good: Driving Social Impact | 19 November 2025 | National University of Singapore

The Economics of Information Pollution in the Age of AI: General Equilibrium, Welfare, and Policy Design

Yukun Zhang，The Chinese University Of Hongkong；Tianyang Zhang，University of Macau

Abstract

- This paper develops a novel general equilibrium model to analyze how generative AI acts as an asymmetric technological shock in information markets—substituting for labor in low-quality content production while merely complementing high-quality creation. This asymmetry systematically lowers the cost of producing "lemons" over "peaches," driving the ecosystem toward a "Polluted Information Equilibrium" sustained by three failures: a production externality, platform governance misalignment, and under-provision of verification. To address these failures, we introduce a real-time Information Pollution Index (IPI) and propose a multi-instrument policy portfolio—including Pigouvian taxes, provenance standards, and fiduciary duties. Agent-based simulations confirm that these interventions generate superadditive welfare gains. Ultimately, the paper reveals a paradox: without proper governance, AI-driven progress can reduce social welfare, making institutional design central to the AI era.

Background &Motivation

- Generative AI dramatically lowers the marginal cost of producing low-quality content.
  AI behaves as a **substitute** for low-effort "lemons" ($\sigma_l > 1$), but as a **complement** for high-quality "peaches" ($\sigma_h < 1$).
  This asymmetric technological shock reshapes information markets:
- Low-quality content becomes extremely cheap and scalable.
- Platforms amplify engagement rather than accuracy.
- Verification is costly and under-provided.
- Trust becomes an eroding **information commons**.
- **Result:** A self-reinforcing **Polluted Information Equilibrium**.

2. Research Questions

How does AI change production costs for low- vs. high-quality information?
Why does the ecosystem converge to high pollution even when actors are rational?
How can we measure system health through a welfare-linked Index?
What policy portfolio can restore efficiency and reduce pollution

Theory Framework and Information Pollution Index(IPI)

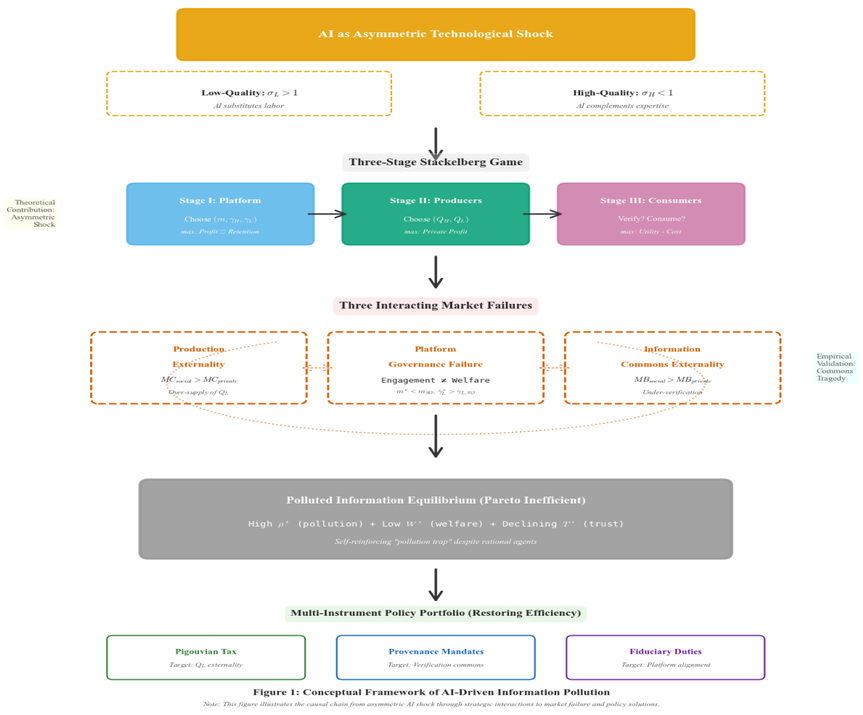
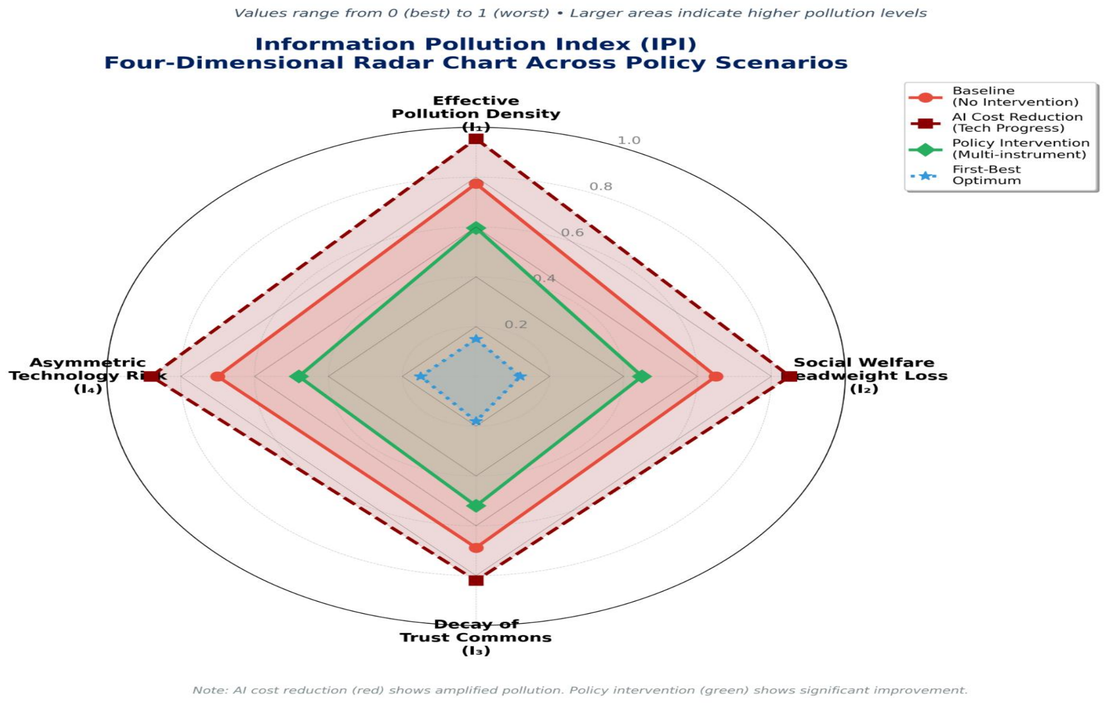


Figure 1: Conceptual Framework of AI-Driven Information Pollution

Note: This figure illustrates the causal chain from asymmetric AI shock through strategic interactions to market failure and policy solutions.

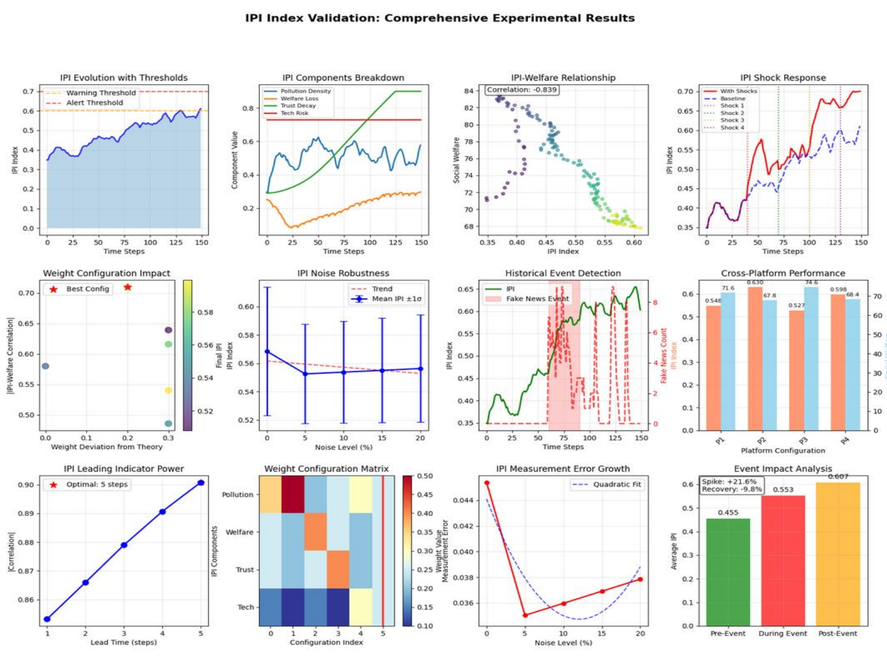


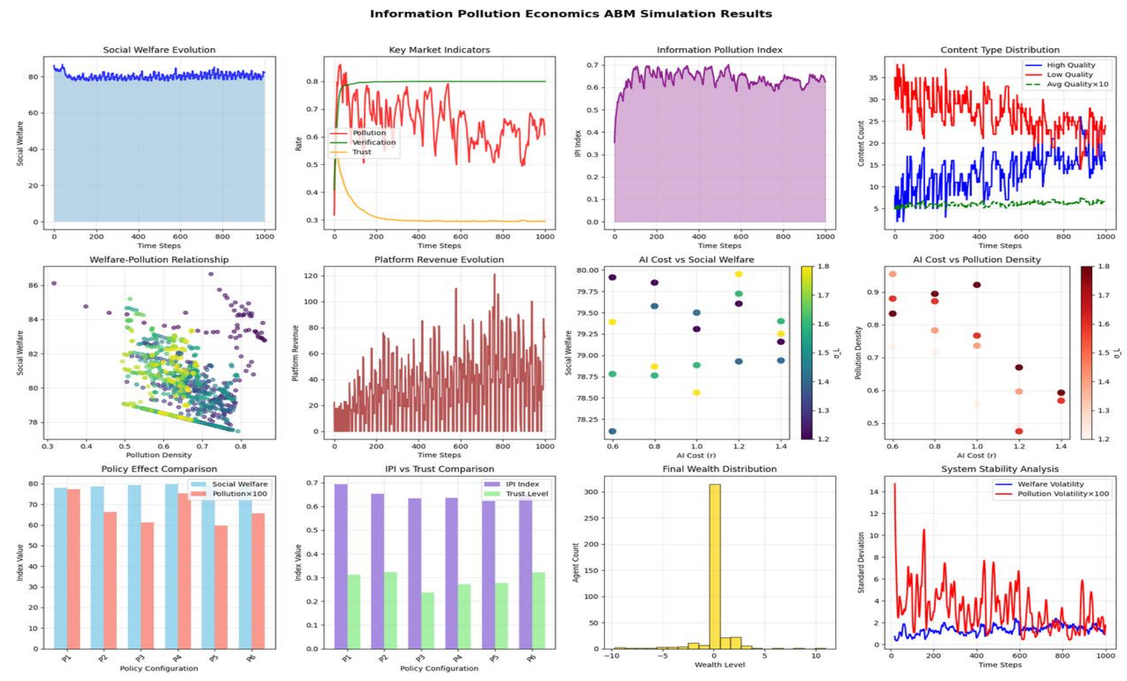
Note: AI cost reduction (red) shows amplified pollution. Policy intervention (green) shows significant improvement.

- Our theoretical framework models generative AI as an asymmetric production shock that fundamentally alters the economics of information creation. In low-quality content production, AI acts as a substitute for human labor ($\sigma_l > 1$), causing costs to collapse, whereas in high-quality content, it functions only as a complement to expertise ($\sigma_h < 1$), yielding far smaller productivity gains. Using a CES production structure, we show that this asymmetry generates a widening cost differential: the marginal cost of "lemons" falls much faster than that of "peaches," increasing the relative profitability of low-quality content as AI improves. Embedding this mechanism into a three-stage Stackelberg model of platforms, producers, and consumers, we prove the emergence of a stable yet Pareto-inefficient Polluted Information Equilibrium, sustained by a production externality, platform amplification bias, and chronic under-provision of verification as a trust commons. Positive feedback loops—expanding low-quality supply, reduced verification incentives, and algorithmic amplification—drive persistent ecosystem drift toward higher pollution levels. The core insight is that information pollution arises not from individual irrationality but from structural incentives shaped by asymmetric AI technology and platform economics.
- The Information Pollution Index (IPI) is a welfare-grounded, multi-dimensional metric designed to measure the overall health of an AI-driven information ecosystem. It captures four structural forces that jointly determine pollution levels: the share of attention captured by low-quality content (Effective Pollution Density), the welfare loss arising from misallocation of consumption and production (Deadweight Loss), the long-run erosion of collective trust due to under-verification (Trust Commons Decay), and the technological asymmetry between generative and detection capabilities (Asymmetric Technology Risk). By aggregating these components into a single continuous index, the IPI provides a real-time "dashboard" for diagnosing systemic drift, evaluating policy interventions, and guiding adaptive regulation in rapidly evolving information markets.

Methodology &Key Finding

- We validate our theoretical model using a bespoke Agent-Based Model (ABM) that simulates heterogeneous producers, consumers, and platform algorithms interacting under varying AI cost levels. The ABM embeds the asymmetric CES cost structure, endogenous verification decisions, and amplification dynamics, allowing us to observe out-of-equilibrium behavioral patterns. Using this environment, we construct and evaluate the Information Pollution Index (IPI), showing that it tracks ecosystem health with high fidelity: the correlation between IPI and welfare is **–0.839**, and the index responds strongly and smoothly to shocks such as reductions in AI cost. The simulations reveal a robust **Paradox of AI Progress**—as the cost of AI declines, low-quality content scales disproportionately, pollution density rises, verification collapses, and overall welfare falls, even though technology is improving. Policy experiments further show that single interventions have limited effect, whereas a multi-instrument portfolio (Pigouvian tax + provenance standards + fiduciary duties) generates **superadditive welfare gains** and achieves the largest reduction in pollution. These findings confirm that both the polluted equilibrium and effective governance are emergent properties of system-level interactions, not individual behaviors.





Policy Implication And Conlusion

- Effective governance of AI-driven information markets requires a coordinated policy portfolio rather than isolated fixes. A **Pigouvian tax** corrects the production externality behind low-quality content, **provenance standards** strengthen verification, and **fiduciary duties** constrain platform amplification incentives. These tools are integrated through the **Information Pollution Index (IPI)**, which functions as a real-time dashboard for ecosystem health and enables **adaptive regulation**—policies automatically adjust when the system drifts toward higher pollution. Simulations show that this multi-instrument strategy generates **far stronger welfare gains** than any single intervention. Ultimately, the results reveal that AI progress can become **anti-welfare** without proper governance, underscoring that well-designed market rules—not technological optimism—determine whether AI improves or harms societal welfare.

Jointly organised by NUS National University of Singapore | Centre for Governance and Sustainability, NUS Business School | Tencent | Supported by CANGO China Association for NGO Cooperation 中国国际民间组织合作促进会