Discussion of "The impact of AI and digital platforms on the information ecosystem" by Joseph E. Stiglitz and Màxim Ventura-Bolet

Wei Li

Vancouver School of Economics, UBC

September 2025

I enjoyed reading and discussing this paper by Stiglitz and Ventura-Bolet. It develops a clean conceptual model of AI's effects on information supply and credibility. Building on Grossman and Stiglitz (1980) and Radner and Stiglitz (1984), it emphasizes that knowledge production and dissemination are forward-looking, costly, and shaped by incentives. The framework contrasts AI-induced gains in aggregation and transmission with losses on the production side, allowing the composition of information to shift through changes in the truth-lie mix and credibility. It also points to policy levers, including adapting intellectual property rules to AI and strengthening accountability. Three points to take away: (i) AI intermediation can erode incentives even as dissemination improves; (ii) credibility depends on lie detection probability relative to the cost advantage of AI-enabled lies; and (iii) policy works through liability, provenance and attribution, and reputational costs.

To sharpen the paper while preserving tractability, I suggest three complementary steps. First, derive comparative statics that clearly separate readership-expansion from business-stealing forces, ideally yielding a threshold or nonmonotonic characterization of when information quantity and credibility rise or fall. Second, the information environment features both misinformation and strategic disinformation operating in parallel: AI lowers the marginal cost of low-quality content, raises the realism and persuasiveness of lies, and enables microtargeting, so targeting itself becomes a driver of the truth-lie composition and credibility. Third, connect the mechanisms to a small set of policy levers by showing how liability and reputation raise the effective cost of lies, and how provenance and attribution reduce uncompensated diversion of attention. Taken together, these steps keep the model disciplined, yield portable, testable implications that can be used for subsequent theoretical and empirical work.

Summary. The model features producers choosing flows of truthful and deceptive content, $I^j(t)$, which accumulate into stocks $Q^j(t)$ that depreciate at rate δ, with $j \in \{T,L\}$ indicating truth (T) or lie (L). Consumers allocate attention according to returns to attention v_j (which depend on the relevant stock of information and a transmission efficiency parameter γ). Digital platforms and AI act through two primary channels: they extract a fraction λ of attention from producers (so producers retain share $1-\lambda$), and they alter transmission efficiency γ. A higher λ reduces the time, and hence revenue, captured by original outlets, weakening incentives to supply new content. The authors evaluate welfare via a function $q(Q^T(t), \gamma)$ (the size/quality of a personal information set), so welfare need not move one-for-one with the information stock.

Two regimes arise. When the fraction of informed consumers (or lie detection probability ω) is high and lying is costly (high θ), the value of lies for uninformed audiences is limited (low vU or low effective exposure), so the equilibrium is *truthful*. Otherwise, an *interior truth-lie equilibrium* obtains. AI can move the system between these regimes by shifting ω (the informed share), θ (relative costs of lies), and returns to attention (v_I, v_U) through transmission technology.

This framing highlights potential “tipping” behavior rather than smooth responses, including the possibility of an “information collapse” when AI becomes sufficiently reliable and λ remains high.

A central implication follows. Even if AI improves transmission efficiency (γ rises), the extraction channel (higher λ) can reduce producers’ captured attention sufficiently to reduce new information supply and the long-run information stock - a modern echo of the Grossman-Stiglitz paradox. In the interior regime, if AI reduces the cost or raises the persuasiveness of lies more than it does for truthful content, credibility declines and polarization intensifies. This can be studied using equilibrium and identification tools in economics, detection and provenance methods in computer science, liability and attribution design in policy, and ranking and interface experiments in design.

In my view, here are some areas the paper may explore more deeply for more theoretical insight and policy relevance.

Net effect on information production. The current analysis is also well suited to study large information aggregators such as Google, and the potential harm has been argued in courts (e.g., recent cases in the EU and Canada).[1] AI is more powerful than traditional aggregators, so both increase transmission efficiency and divert attention, with AI doing so more strongly. If AI and information aggregators improve returns to attention (v_I and v_U) but also raise extraction (higher λ), can the authors state a threshold that characterizes the net effect? This can lead to testable comparative statics results on the trade-off between business-stealing and readership

[1] Both the EU’s Copyright Directive (Article 15) and Canada's Online News Act (Bill C-18, 2023) require platforms like Google and Meta to compensate news publishers for journalistic content. In Canada, regulators have imposed a compliance fee on Google to cover enforcement costs of the Online News Act, which mandates platform funding for news content. See https://www.reuters.com/technology/canada-regulator-impose-fee-google-online-news-laws-operating-costs-2025-02-27/, Reuters, Feb 27 2025.

expansion effects of news aggregators (e.g., Jeon and Nasr 2016). It also helps interpret the empirical evidence offered by Athey, Möbius and Pál (2021).

Knowledge markets vs. news markets. The paper focuses on news providers (traditional media and online outlets), which can differ from the broader knowledge production sector. In news, information is financed largely through attention, so it behaves more like a public good; in R&D and science, appropriability (IP, licensing, subscriptions, grants) provides private incentives. It would help to distinguish these cases more sharply, especially for depreciation (news obsolescence vs. scientific irrelevance) and for policy mapping, while asking whether the difference is one of *kind* or primarily of *degree*. In particular, if AI becomes the default interface to knowledge (answer engines, tool-integrated assistants) and weakens appropriability or reroutes usage away from originators, the effective exposure to intermediation can converge toward the news case. A brief statement of conditions under which the comparative statics (e.g., the sensitivity of Q^j to changes in λ) coincide across the two markets (for example, when effective λ and depreciation profiles are similar) versus when they diverge would clarify where concerns remain modest and where they may become first-order.

Misinformation and strategic disinformation. AI changes the problem qualitatively as well: it lowers the marginal cost of high-fidelity synthesis (text, audio, video), making lies more realistic and harder to detect. It also enables microtargeting of susceptible audiences, including those who would otherwise count as informed. Because attention is a common pool, such misinformation is effective and dilutes the usefulness of all information. Li and Tan (2025) show that personalized targeting *endogenously generates polarization*, with stronger strategic influencers focusing on opinion leaders and leaving users in perpetual disagreement. Empirical evidence also points to a deterioration in information quality: newsroom production costs and workflows shift with AI (Sonni et al. 2024); platform algorithms foster filter bubbles and fake news (Mohseni and Ragan 2018); generative systems can propagate factual errors and reduce diversity in reporting (Brantner, Karlsson and Kuai 2025); firm-level misinformation measurably moves markets (Fan et al. 2024); and probabilistic sharing improves lie detection relative to binary judgments (Guilbeault, Woolley and Becker 2020).

The framework can accommodate such mis- and disinformation in three ways. First, it can model continuous quality and probabilistic detection. Replace the binary truth-lie quality with a continuous quality index q, and let detection be $d(q,\omega)$, decreasing in q and increasing in the fraction of informed users ω. This captures AI's realism and persuasiveness upgrade for lies (deepfake realism) and yields testable comparative statics. Second, it can model endogenous informed share. Make the share of informed users the outcome of costly verification (time, tools, provenance), allowing AI to shift both the cost of becoming informed and the noise of verification (hallucinations). This delivers a clean threshold linking screening ω, verification cost θ, and the deception cost gap. Third, the model can allow for targeted persuasion by allowing nonproportional targeting of the uninformed (opinion leaders, high-centrality nodes), so senders choose whom to reach, not just how much to produce, mapping directly to strategic competition among influencers (Li and Tan 2025).

Welfare and policy implications. The model supports meaningful welfare analysis. In particular, what is the socially optimal level of platform intermediation λ? Should policy subsidize producers, regulate access, or tax intermediaries? A planner's problem from a representative consumer's perspective could serve as a useful benchmark. Yet in the absence of calibrated simulations or a richer equilibrium characterization, the reader is left unsure how large the effects might be or which channels matter most. Even stylized calibrations (with back-of-theenvelope parameters for $\lambda,\omega,\theta,\gamma$) would indicate whether realistic policy shifts can restore the truthful regime. A practical way to implement this distinction is to treat *validated (v) and unvalidated (u) channels* as separate policy targets. For validated outlets, policies should increase detection and reduce diversion (higher ω, lower λ_v): think attested provenance, stronger attribution, and revenue-sharing/traffic routing that preserves incentives to originate. For unvalidated channels, policies should raise the expected cost of deception (higher θ_u) via liability, reputational penalties, and distribution limits for unverifiable claims. This separation gives regulators and platforms two orthogonal levers: move traffic toward the validated sector and raise the private cost of deception outside it. Note that today's hallucinations partly constrain substitution away from sources; as AI accuracy improves, that brake weakens, making the λ-γ trade-off steeper.

Looking ahead, this paper provides a clear and timely framework for analyzing how advances in AI and platforms reshape the incentives to produce information. By clarifying threshold conditions, separating targeting from quality, and adding a compact welfare benchmark, the framework can serve as a foundational tool for understanding and regulating the information ecosystem in the AI era.

References

- Grossman, Sanford J., and Joseph E. Stiglitz. 1980. "On the impossibility of informationally efficient markets." American Economic Review 70 (3): 393–408.
- Radner, Roy, and Joseph E. Stiglitz. 1984. "A nonconcavity in the value of information." In Bayesian Models in Economic Theory, edited by Marcel Boyer and Richard E. Kihlstrom, 33–52. Amsterdam: Elsevier.

- Jeon, Doh-Shin, and Nikrooz Nasr. 2016. "News aggregators and competition among newspapers on the internet." American Economic Journal: Microeconomics 8 (4): 91–114.
- Athey, Susan, Markus Mobius, and Jeno Pal. 2021. "The impact of aggregators on internet news consumption." NBER Working Paper 28746. Cambridge, MA: National Bureau of Economic Research. https://www.nber.org/papers/w28746.
- Sonni, Alem Febri, Hasdiyanto Hafied, Irwanto Irwanto, and Rido Latuheru. 2024. "Digital Newsroom Transformation: A Systematic Review of the Impact of Artificial Intelligence on Journalistic Practices, News Narratives, and Ethical Challenges." Journalism and Media 5 (4): 1554–1570.
- Mohseni, Sina, and Eric D. Ragan. 2018. "Combating fake news with interpretable news feed algorithms." arXiv preprint arXiv:1811.12349. https://arxiv.org/abs/1811.12349
- Brantner, Cornelia, Michael Karlsson, and Joanne Kuai. 2025. "Sourcing behavior and the role of news media in AI-powered search engines in the digital media ecosystem: Comparing political news retrieval across five languages." Telecommunications Policy 49 (5): 102952.
- Fan, Jianqing, Qingfu Liu, Yang Song, and Zilu Wang. 2024. "Measuring misinformation in financial markets." SSRN Working Paper. https://ssrn.com/abstract=4922648.
- Li, Wei, and Xu Tan. 2025. "Dynamic network influence: The art of strategic messaging." SSRN Working Paper 5066467. https://ssrn.com/abstract=5066467.
- Guilbeault, Douglas, Samuel Woolley, and Joshua Becker. 2020. "Probabilistic social learning improves the public's judgments of news veracity." arXiv preprint arXiv:2010.06019. https://arxiv.org/abs/2010.06019.