Economists can help ensure the AI disruption brings society to the good place: discussion for "The Coasean Singularity? Demand, Supply, and Market Design with AI Agents"

David Rothschild

"The Coasean Singularity? Demand, Supply, and Market Design with AI Agents" (CSDSMDAI) presents a rich and ambitious agenda, touching on a wide array of themes at the intersection of artificial intelligence (AI) and economic design. At a high level, the authors explore how AI agents can participate in economic systems focusing on consumer demand, product design and supply, and the formation of markets and equilibrium.

The breadth of the CSDSMDAI spans multiple core economic concepts while simultaneously addressing an extensive list of practical and ethical concerns. These include preference elicitation and alignment between agents and their principals, identity verification, and important societal challenges such as bias, privacy, and consumer protection.

Given the scope and complexity of the issues raised, this discussion will selectively focus on a few central themes. To maintain clarity, I will use the term *principal* to refer generically to any of the three key stakeholders that deploy AI agents, defined as:

- **Consumers**: People who want goods or service may be represented by *assistant agents*.
- **Businesses**: Firms that provide goods or services and may deploy *service agents* to interact with consumers or their assistant agents.

- **Platforms**: Large intermediaries or marketplaces such as Amazon, Expedia, OpenTable, or Spotify that host and facilitate interactions among many consumers and businesses, increasingly deploying their own platform agents.

Stages of Disruption

I categorize the disruption introduced by AI into three somewhat inexact stages.

1. **Augmentation within existing workflows:** In this initial stage, principals (i.e., consumers and businesses) employ AI agents to support or enhance specific tasks within their current workflows. For example, a consumer might delegate a detailed search for a durable good to an assistant agent, while still retaining control over the subsequent stages of the conversion journey. Much of the earliest economics research focused on the productivity gains within this stage (e.g., Brynjolfsson et al. 2025, Cui et al. 2024, Spatharioti et al. 2025).
2. **Redistribution of effort within workflows:** As AI capabilities mature, principals begin to reallocate their time within the same workflows. Tasks that once required significant manual effort can now be executed more efficiently, freeing time for new activities. For instance, if an agent handles email triage and drafts presentation materials, a salesperson might shift more attention to direct client engagement. This has proven more difficult to research, as it really requires both adoption and time in the field to materialize, but work is starting to materialize (e.g., Dillon et al. 2025).
3. **Reconstruction of workflows:** In the most advanced stage, workflows are re-designed from the ground up to fully leverage the capabilities of AI agents. At this point, agent integration is not simply a matter of efficiency, it redefines the structure of economic activity itself. Obviously, at this point, the work needs to be theoretical or building out the baselines (and monitoring) necessary to understand

these disruptions when they occur (e.g., Rothschild et al. 2025)

In CSDSMDAI's discussion of demand, the authors argue that high-stakes processes, particularly those already mediated by human agents, are the most fertile ground for early adoption of AI agents. Table 1 identifies tasks with the greatest potential for augmentation, emphasizing those that require extensive information gathering, review, or monitoring. These tasks align well with AI's comparative advantages over humans, who are limited by cognitive bandwidth and time. This framing is compelling within a Stage One or Stage Two paradigm, where the investment in agentic infrastructure is justified by the value of augmenting existing workflows, and principals are well-positioned to deploy agents. However, once systems transition into Stage Three, the calculus shifts. At that point, returns may emerge from low-stakes but high-volume transactions that are infeasible with humans- in-the-loop. For instance, micro-content or other unbundled goods and services designed for machine consumption at sub-cent price points could thrive in markets without user interfaces, where transaction costs render human participation impractical. Moreover, many businesses that currently operate through traditional websites may evolve toward an *agent-first* model, fundamentally altering the boundary between tasks best suited for agents versus those requiring human involvement.

Incentives and Monetization

In CSDSMDAI, the authors distinguish between "bring-your-own-agent" and "bowling shoe agents," which align closely with what we describe in Rothschild et al. (2025) as consumers' *assistant agents* (currently commonly in the form of *computer-use agents*) and businesses' *service agents* currently commonly in the form of *siloed agents*, respectively.

Computer-use agents function by mimicking human interactions with existing web-based interfaces—often scraping content, navigating forms, and simulating clicks. This is

noisy, and as the authors note, many businesses are already taking steps to restrict such behavior because it is inherently adversarial to their agency and business model. As such, we view this as an unsustainable model for broad agentic integration.

Siloed agents represent conversational front ends appending to existing human interfaces, for instance a chatbot interface acting as a sales representative for a business on their website. While these agents are more stable and aligned with business interests, they serve as a transitional phase. They offer first-stage benefits by improving access and responsiveness, but are clearly incentively-aligned with the business, not the consumer they purport to represent.

This is why we argue in Rothschild et al. (2015) that the market is likely to evolve toward a *bring-your-own-agent* paradigm for consumers, not in the form of today's computer-use agents, but as consumer-controlled assistant agents designed to interface directly with businesses' service agents and APIs. These agents will be optimized for agentic ecosystems rather than legacy web environments, reducing noise and misaligned incentives, while creating a richer information environment. As a result, many of the dynamics discussed in CSDSMDAI—particularly those related to market efficiency, negotiation, and consumer choice—are likely to shift substantially. Further, while CSDSMDAI focuses primarily on the consumer side, examining demand, supply, and market dynamics, the evolution of business-side service agents and the intermediary marketplace is equally central to understanding the broader development of consumer assistant agents.

With these assumptions many of the questions presented in Table 2 of CSDSMDAI take on a different dimension. For example, as we transition from an *attention economy* to a *preference economy* (Rothschild et al. 2025), the nature of advertising will evolve. In the pessimistic (bad place) scenario, businesses use ads to get preferred access that is not correlated with the consumer's preference. In the optimistic (good place) scenario,

instead of simply attempting to capture a principal's (or even their agent's) attention through ads, businesses will redirect their advertising spend to create products of higher quality that better align with consumer preferences, more informed and useful agents, and to validate their quality.

Centralized platforms currently dominate digital time and spending due to their advantages in data aggregation, processing power, and reduced communication friction (further advantages in discovery, validation, remediation, and economies of scale in production and distribution build on this moat). In Rothschild et al. (2025), we speculate that an agentic economy would empower consumers and businesses with greater control over data storage and sharing, and access to processing power. This shifts the critical bottleneck to communication friction. While natural language agent-to-agent communication could, in theory, reduce this friction to near zero—enabling seamless interaction across services—by artificially restricting agent communication within their ecosystems, they can maintain high switching costs and lock-in effects, preserving their market power. If platforms succeed in sustaining high communication costs through interoperability constraints, they will continue to capture a disproportionate share of welfare surplus and stymie innovation, limiting the realization of CSDSMDAI's projected benefits.

Walled gardens remain the status quo because they offer lower risk for key stakeholders by controlling access and minimizing downside threats. However, a more open marketplace would promote more equitable welfare distribution and greater innovation. Realizing a truly open web of agents will require not only technical infrastructure but also socio-technical and social science frameworks to ensure the quality, trust, and safety necessary for consumers and businesses to move beyond the relative ease and security of walled gardens.

Agent Performance

CSDSMDAI offers a thoughtful and measured discussion of assistant agents. However, one important point, implicit in CSDSMDAI, deserves more explicit attention. Much of the literature evaluating the reliability or quality of AI agents tends to compare them not to the humans (or human operated systems) they are intended to augment or replace, but to idealized, near-omniscient benchmarks. This mirrors a pattern seen in early work on algorithmic bias, where algorithmic decisions were often judged against unrealistic standards rather than the flawed human judgments they were replacing.

This misalignment in baseline comparisons is particularly consequential in tasks involving preference elicitation. An agent's inability to accurately infer (and then execute on) a principal's preferences is often not a failure of the agent, but a reflection of the principal's own imprecise, latent, or evolving understanding of their preferences. In such cases, the agent's role may be less about precision and more about helping the principal refine, articulate, and discover what they actually want. Early research suggests AI agents may be especially effective in this domain (Droso et al. 2025).

Closely related is a distinction the authors acknowledge but do not fully explore: the difference between an assistant agent *assisting* in a task versus *transacting* on behalf of the principal. For example, when an agent conducts a detailed search for a durable good and presents curated options, its performance should be evaluated against what the consumer could have achieved independently in terms of time, research depth, and quality. However, if the agent is empowered to complete the transaction—potentially involving negotiation, commitment, or significant financial stakes—the complexity and consequences of misalignment increase substantially. The design and evaluation of agentic systems must account for these contextual distinctions and apply appropriate human baselines for comparison.

These comparative frameworks are essential when considering the efficiencies,

entitlements, and protections required for agents and marketplaces to support a more open web-of-agents.

Discussion

Economists have a critical role to play in shaping how AI agents will disrupt economic systems in the coming years. The welfare implications are profound, particularly in determining whether agentic ecosystems evolve into fragmented walled gardens or open, interoperable marketplaces. Economic research can help design mechanisms that ensure agent alignment with their principals, minimize fraud, reduce bias, protect privacy, and safeguard both consumers and businesses making a web-of-agents viable.

Beyond technical design, economists are uniquely positioned to influence the regulatory frameworks, institutional structures, and societal norms that will govern agent-driven markets. Their contributions will be essential in ensuring that the welfare gains from these technologies are distributed equitably and that agentic systems serve the broad public interest, not just the most powerful incumbents. In short, economists can help ensure the AI disruption brings society to the good place.

References

[1] Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond. Generative ai at work. *The Quarterly Journal of Economics*, page qjae044, 2025.

[2] Cui, Zheyuan Kevin, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz. The effects of generative ai on high skilled work: Evidence from three field experiments with software developers. *Available at SSRN 4945566*, 2024.

[3] Dillon, Eleanor W., Sonia Jaffe, Nicole Immorlica, and Christopher T Stanton. Shifting work patterns with generative ai. Technical report, National Bureau of Economic Research, 2025.

[4] Drosos, Ian, Jack Williams, Advait Sarkar, Nicholas Wilson, Sean Rintel, and Payod Panda. Dynamic prompt middleware: Contextual prompt refinement controls for comprehension tasks. In *Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work*, pages 1–23, 2025.

[5] Rothschild, David M., Markus Mobius, Jake M Hofman, Eleanor W Dillon, Daniel G Goldstein, Nicole Immorlica, Sonia Jaffe, Brendan Lucier, Aleksandrs Slivkins, and Matthew Vogel. The agentic economy. *arXiv preprint arXiv:2505.15799*, 2025.

[6] Spatharioti, Sofia Eleni, David Rothschild, Daniel G Goldstein, and Jake M Hofman. Effects of llm-based search on decision making: Speed, accuracy, and overreliance. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2025.