The Coasean Singularity? Demand, Supply, and Market Design with AI Agents

Peyman Shahidi
MIT

Gili Rusak
Harvard

Benjamin S. Manning
MIT

Andrey Fradkin
BU & MIT IDE

John J. Horton*
MIT & NBER

September 16, 2025

Abstract

This paper examines the transformative economic implications of AI agents—autonomous software systems that can perceive, reason, and act in digital environments to achieve goals on behalf of human principals. We focus on their role as market participants, capable of executing tasks such as search, negotiation, and communication, and of transacting directly in digital marketplaces. Demand for AI agents is a form of derived demand, that will be shaped by agent quality, industry context, and users' trade-offs between decision quality and effort reduction. On the supply side, firms will compete to design, deploy, and monetize agents, and how this competition interacts with platform governance. At the market level, we highlight both efficiency gains through reduced transaction costs and efficiency losses through increased frictions. AI agents expand the frontier of feasible market design by reducing the costs of preference elicitation, contract enforcement, and identity verification, while simultaneously raising novel regulatory challenges around bias, privacy, and consumer protection. These factors create many opportunities for impactful economics research.

* Emails: Shahidi: peymansh@mit.edu; Rusak: gilirusak@g.harvard.edu; Manning: bmanning@mit.edu; Fradkin: fradkin@bu.edu; Horton: jjhorton@mit.edu.

1. Introduction

AI agents are autonomous software systems that perceive, reason, and act in digital environments to achieve goals on behalf of principals, with capabilities for tool use, economic transactions, and strategic interaction.[1] This paper focuses on the transformative implications of these systems as market participants. We envision AI agents as having the ability to harness computational resources, to communicate with other agents and humans, to receive and send money, and to access and interact with the Internet.[2] A variety of AI agents are already available, and more capable ones are in the development pipeline.

Perhaps the prototypical AI agent is Deep Research (Citron, 2024), introduced by Google's Gemini team and quickly copied by other players including OpenAI and Anthropic. Current versions of Deep Research take natural language instructions (prompts) provided by a human, and autonomously search, reason, and use tools to generate a researched report in response to the instructions. Deep Research allows us to illustrate the distinction between agents and traditional software. An economist writing a paper without Deep Research might use Google Scholar to find papers, refine queries, and synthesize results—software retrieves information but leaves reasoning to the user. Alternatively, they could upload papers into an AI system and ask it to generate a report, which draws only on the provided data.[3] In neither case does the software autonomously define and pursue a course of action. In contrast, Deep Research can iteratively search the web, evaluate results, and generate a final report. This autonomy—perceiving, reasoning, and acting in digital environments to achieve goals conveyed via natural language—is what makes AI agents distinctive.

We provide a practical, consumer-oriented perspective on AI agents, complementing the other handbook chapter in this research agenda (Hadfield and Koh, 2025), which provides a theoretical perspective on AI agents. We take seriously the idea that rather than hiring a human agent, coach, matchmaker, negotiator, or salesperson, one could use an AI agent instead. This possibility is near, as AI agents can already imperfectly execute many tasks related to search, preference elicitation, negotiation, and communication (e.g. Eloundou et al., 2024, Allouah et al., 2025, Zhu et al., 2025, Dobbala & Lingolu, 2024, Dammu et al., 2025). In fact, AI agents can already search for products on e-commerce websites (e.g. Google Project Mariner, Zeff, 2025, Allouah et al., 2025). One could easily imagine more advanced AI agents that could contact multiple car dealerships, and negotiate, or to apply to jobs and advocate for getting hired on the user's behalf (Immorlica et al., 2024).

The fundamental economic promise of AI agents lies in their ability to dramatically reduce transaction costs—the expenses associated with using markets to coordinate economic activity. This reduction can occur through two primary mechanisms: direct task execution and market facilitation. In the direct approach, agents perform tasks entirely on behalf of users, from price comparison shopping to contract negotiation to job interviews. In the facilitation approach, agents assist users in making better market decisions, such as helping job applicants individually optimize their resumes

[1] We use the term "principal" to mean any stakeholder that deploys AI agents. The principal can be a consumer or firm that hires an assistant agent to act on their behalf in a market activity, or a business/platform that deploys service agents to interact with consumers or with consumers' representative agents.
[2] For the purposes of this essay, we abstract away from Artificial Super Intelligence (ASI), in which AI is omniscient in every dimension relative to humans.
[3] Currently, the term 'Large Language Model' is commonly used to describe AI models with advanced capabilities. However, we avoid using this term throughout the essay because frontier AI architectures and their popular names will likely evolve over time. Instead, we opt for "AI agent" or "AI system"

for each job submission. Recent evidence suggests both approaches can be quite effective. For the former, one large study has shown that outcomes can be *better* for job seekers interviewed by AI as opposed to human employers (Jabarian & Henkel., 2025). For the latter, algorithmic assistance on job application essays has been shown to increase hiring rates significantly (Wiles et al., 2025). However, this automation also introduces new complexities: widespread AI assistance can obscure genuine human effort and homogenize market-facing communications—applications, proposals, listings, reviews, and negotiation messages alike.

It is important to recognize that demand for AI agents represents derived demand rather than direct utility. Individuals do not generally derive satisfaction from watching an agent compile price lists for gas grills; they employ the agent purely to achieve some desired market outcome following their own decision-making process. As a result, humans will deploy AI agents in two primary scenarios: first, to optimize decisions that they would otherwise make sub-optimally due to information constraints or cognitive limitations; and second, to make decisions of similar or potentially even lower quality, but at dramatically reduced cost and effort.

As agents proliferate, there will likely be transformative downstream effects. This potential stems from the same fundamental insight that makes Coase's transaction cost economics so influential in understanding economic organization (Coase, 1937). One could argue that much of how we structure our economy and firms can be explained by transaction costs, often costs of human labor. The activities that comprise transaction costs—learning prices, negotiating terms, writing contracts, and monitoring compliance—are precisely the types of tasks that AI agents can potentially perform at very low marginal cost. If agents can indeed execute these functions effectively and cheaply, we may witness significant shifts in the traditional make-or-buy boundaries that define firm organization and market structure. Of course, similar technologies may also affect the costs of producing goods and services, which is an issue beyond the scope of this article.

We expect these transformations to unfold in three stages: (1) deployment of AI agents within existing workflows—principals deploy agents to support or enhance discrete tasks while the process and decision rights remain unchanged; (2) redistribution of effort within workflows—as capabilities mature, agents become better at completing routine steps and human time is reallocated toward judgment, relationship work, and oversight (or the same workflows are executed with higher frequency); and ultimately (3) reconstruction of workflows—processes are redesigned around novel agent capabilities (e.g., agent-to-agent interfaces). Broadly, each stage substitutes AI for human effort, progressively pushing markets toward agent-first structures in the long-run.

The spread of AI agents is not guaranteed to make markets more efficient, however, or to improve other social objectives. Even if it's individually rational for consumers and firms to adopt these agents in their workflows, the equilibrium outcome may be suboptimal. For instance, the dramatically lower cost of submitting job applications has led to excessive congestion in the labor market. Many firms now report receiving several times more applications than in previous years (Kessler, 2025).

At the same time, widespread adoption of agents expands the economic market design frontier. By collapsing the costs of eliciting preferences, enforcing commitments, and verifying identity, they make mechanisms that were once only theoretically attractive now feasible at scale. The upshot is not just faster versions of the status quo designs, but rather a larger feasible set of design options. This points to an exciting market design agenda for economists, who are uniquely well positioned to guide this transition by translating theoretical insights into practical designs that can harness the full benefits of AI agents.

The rest of the article proceeds as follows. In Section 2, we discuss demand for AI agents, and how it will vary depending on agent quality and industry. In Section 3, we discuss the economic design of AI agents, and the relationship of this design to the alignment problem. In Section 4, we consider the supply of agents, and how companies will compete in offering AI agents to consumers. In Section 5, we discuss the equilibrium implications of the widespread adoption of AI agents. In Sections 6 and 7, we consider market design and regulatory issues related to AI agents, respectively, and finally we conclude in Section 8.

2. Demand for AI Agents

Humans will demand AI agents for fundamentally the same reasons they demand human agents: they believe it is rational to have some aspect of a decision or market transaction handled through an intermediary. This might occur because the agent's time is less expensive than the principal's time, because the agent is legitimately better at the task than the principal, or because the principal has reasons for obscuring their identity in the transaction.

These motivations will increasingly drive the adoption of AI agents in two avenues. First, agents will substitute for human intermediaries where one would otherwise either do the work personally or hire a human agent. The canonical example is product search. AI agents convert the costly, time-consuming parts of intermediation—search, screening, quoting, negotiation, scheduling—into low-cost compute and API calls. An agent can solicit and compare many quotes in parallel, then book and monitor follow-through at far lower marginal cost than a human. Even as skilled execution remains human, the intermediation premium falls, creating demand for AI agents.

Second, and perhaps more consequential, agents will make viable many tasks one would not have attempted at all. By pushing the cost of exploration and execution down, they expand one's feasible set and lower the threshold of “worth doing.” For physical jobs (e.g., home installation, fixing a sink), the agent can conduct diagnostic triage, source parts, select vendors, and schedule service whereas for software, it can generate and iterate on a bespoke script (Chatterji et al., 2025). With low marginal costs, agents can persist through repeated failures to eventual success, and continue the monitoring tasks over longer windows to achieve a more desirable outcome.This same cost dynamic will also enable new forms of market engagement that are currently infeasible. We discuss these in Sections 3 and 6.3.

Intertwined with demand will be evaluation. And with AI agents, we will want the same core attributes that we seek in human agents: capability sufficient to act on our preferences successfully, knowledge of our preferences, and alignment sufficient to act on our preferences to our benefit. In essence, we want capable, knowledgeable, and faithful agents. There are substantial theoretical and practical challenges in ensuring that agents have these characteristics.

Although the reasons for the demand for AI agency are clear, how this will translate into a willingness to pay is less clear. We do not expect that consumers will demand perfect AI agents, just as they don't demand perfect human agents. But they will want to know the boundary conditions when AI agents may fail. Practically, communicating capabilities and failure modes to consumers will be a challenge. Another related challenge is the assignment of blame when AI agents fail. There will be substantial demand for information about agent performance, with benchmarks playing important roles alongside word-of-mouth recommendations as people rightly worry about benchmark gaming. Differentiation will likely emerge, with some agents becoming known as particularly effective for certain applications.

We will likely use agents to help us find and evaluate other agents. Current agents might interview new agents, put them through their paces, or share information as users upgrade to better agents. This bootstrapping process will likely involve reviews, benchmark comparisons, and simulation testing to establish agent capabilities and reliability.

AI agents will gain initial traction in markets where human agency is already common. As listed in Table 1, key characteristics include high-stakes transactions, vast spaces of potential counter-parties, substantial effort required to evaluate counter-parties, information asymmetries that could be resolved through effort or experience, and experience asymmetries due to differences in transaction frequency. Markets fitting these criteria include job search, real estate, and certain investment contexts where parties care about counter-parties beyond purely financial considerations. A useful heuristic suggests AI agents will appear first in markets that already employ human agents or where large digital platforms were created to overcome matching frictions—LinkedIn, Upwork, Zillow, Airbnb.

Market Characteristic	Example Markets	Existing Solutions	How AI Agents Help
High-stakes transactions	Real estate, Job search, Investment decisions	Human agents (realtors, headhunters, financial advisors)	AI agents can analyze vast amounts of data and documentation without fatigue, providing thorough due diligence at near-zero marginal cost.
Vast counter-party space	Dating, Freelance hiring, Rental markets	Digital platforms (Tinder, Upwork, Airbnb)	AI agents can evaluate thousands of options simultaneously, with no opportunity cost to their “time” – they can search exhaustively where humans must sample.

High evaluation effort	Startup funding, College admissions, B2B procurement	Specialized consultants, matching services	AI agents can read every review, analyze every metric, and compare all attributes across options without the time constraints that force humans to use heuristics.
Information asymmetries	Used car markets, Insurance shopping, Legal services	Brokers, comparison sites, expert intermediaries	AI agents can continuously monitor multiple information sources, cross-reference data, and identify discrepancies that would take humans hours to uncover.
Experience asymmetries	Home buying (once per decade vs. daily), Wedding planning, Estate planning	Professional agents who transact frequently	AI agents can leverage collective experience from millions of transactions, effectively giving every user the negotiating power of a frequent transactor.

3. Designing AI Agents

In the previous section, we explored why humans will demand AI agents. But of course, this product/service does not yet exist, at least in the form being imagined. In this section, we consider the design and development of agents. While we cannot hope to describe the precise design, we can speculate on the key challenges and the focus of R&D efforts. The use of AI agents will be driven by their design, which has both an engineering and an economic component. On the engineering side, there are practical challenges in having capable agents that can interact with the digital world in a reliable manner. We set aside the engineering challenges for the purposes of this discussion, though they are substantial in their own right (Kalai et al., 2025). Instead, we focus on the economic component, namely what actions should a capable agent take.

Clearly, an AI agent must know the principal's preferences to act on their behalf. The core design problem is thus two-fold: both eliciting those preferences and ensuring the agent honors them—together these are what computer scientists and philosophers refer to as the alignment problem (Bostrom, 2014, Christian, 2020). Note that this encompasses the economic principal-agent problem, where preferences are already known and the challenge is designing contracts or incentives to prevent self-interested agents from shirking their duties.

Preference elicitation, in particular, offers new challenges even though the practice itself is not new to digital markets. Conventional recommendation systems already translate billions of online consumer choices into predicted preferences using complex machine learning algorithms. While these traditional systems are powerful, their flexibility remains fundamentally limited. They operate with fixed input and output dimensionalities and are trained for specific contexts. For example, Netflix's current television recommendation algorithm cannot recommend consumer or financial goods—even if streaming choices contain latent information about such products.

Flexibility, however, can lead to unexpected outcomes because it is not possible for a principle to fully specify all edge cases an agent might encounter. As such, failures to recover a principal's "full" preferences can still arise for two distinct reasons: (i) principal's articulation limits: the principal cannot fully or consistently specify their preferences (e.g., providing a rank-ordering or pairwise comparisons of the entire Netflix catalog is impractical, even if an individual could in principle generate such an ordering), and (ii) agent's synthesis limits: the agent misinterprets what is stated, including via inaccurate inference or hallucination (e.g., Ji et al., 2023, Manakul et al., 2023, Spatharioti et al., 2025).

As AI models improve, they infer preferences more accurately from the principal's inputs. Accordingly, the portion of the gap between a principal's actual preferences and the agent's representation of it attributable to agent miscomprehension shrinks, while the portion due to principal-side uncertainty or inarticulability may persist. Thus, in the near term, agents function as aggregators of <u>expressed</u> information rather than discoverers of latent tastes; and residual gaps reflect input limits, not inference failures.

Foundation models already offer unprecedented flexibility in learning preferences.[4] They are trained to reason directly over natural language, enabling a single agent to parse a paragraph describing someone's tastes and immediately rank jobs, products, or media options. More importantly, they can make reasonable assumptions when preferences are incompletely specified, drawing on learned relationships from enormous training corpora to fill information gaps.

Consider recent empirical work by Rusak et al. (forthcoming), who demonstrate how agents can flexibly apply human preferences to select jobs. In one study, they asked crowdsourcing workers described in writing the types of tasks they generally like to complete on Prolific.[5] For example, one participant wrote: "I prefer to complete tasks that include games and puzzles. Surveys are fine, if they don't ask me too many personal questions. I never do a task that takes more than 30 minutes, pays less than $10 per hour, or requires me to record audio". An LLM provided with these instructions could easily rank diverse tasks on Prolific's website on behalf of this participant—clearly preferring testing a new version of Wordle over rating cereal logos, which would be preferred over recording audio. The agent could even assess unstated preferences, reasonably inferring that someone who dislikes audio recording would also dislike video recording.

Sophisticated AI agents could go beyond preference inference to facilitate preference discovery. Although this is perhaps underappreciated by economists, humans often don't have a good sense of their own preferences. Similar to how a good coach or therapist helps people discover these

[4] A foundation model is a large-scale AI model trained on broad, diverse datasets that serves as a base for multiple downstream applications.

[5] For those unfamiliar, Prolific is an online platform for recruiting human participants to take part in research studies. Participants browse a feed of available studies or "tasks," which list compensation, time requirements, and other details. They can choose which tasks to complete, and are paid directly through the platform upon completion.

preferences, an AI agent can help people engage in self-discovery by observing patterns in behavior that individuals might not consciously recognize. For instance, an AI home-buying assistant might notice that a buyer consistently gravitates toward houses near parks, with large windows, three bedrooms, and built after 2000—even though the buyer has never explicitly articulated these as priorities. By highlighting such patterns and asking reflective questions, the agent could help individuals articulate values they hold but haven't fully expressed, such as a preference for natural light, newer construction, or access to green space.

Agent-based preference discovery may remain limited, however, as complexity of preference elicitation varies significantly across domains. In some cases, preferences are relatively straightforward: a home seller wants to maximize price while minimizing time to sale, requiring the agent to understand only the speed-price trade-off. In contrast, a buyer's agent needs extensive knowledge like that listed in the example above. When the dimensionality of a principal's preferences far exceeds what they can easily express, performance may deteriorate (Liang, 2025). Consequently, we anticipate that preference dimensionality for a given context will serve as a key predictor of agent adoption patterns.

4. Supply of AI Agents

In this section, we consider the production of AI agents and competition between AI agent providers. Already there are two types of providers: those who build their own foundation models (e.g., Anthropic and OpenAI) and those who use others' foundation models and customize them for particular use cases (e.g., Decagon, Harvey, and Sierra). The economics are different for these two modes, since training foundation models incurs high fixed costs while operating agents incurs mostly variable costs.

We envision that any AI agent provider needs the ability to access and fine-tune a foundation model. It is an open question whether winning AI agent providers need to build their own foundation models. If they do, then we are likely to see a level of concentration that mirrors the concentration in the foundation model industry (Fradkin, 2025). AI agent providers will also need to optimize the usage of the foundation model and to create agent capabilities for interacting with the world. The provider, if it has proprietary data, can let the agent use this data or can fine-tune the agent on this data (Zhang et al., 2024).

In terms of ownership, there will be two types of agents: *bring-your-own* and *bowling-shoe*. Presumably, few consumers will literally create their own models; by "their own" we mean an agent created by some marketplace or platform that is a participant in an exchange. For example, an Anthropic-powered agent being used on Walmart and Amazon via public APIs is a bring-your-own agent: it carries your instructions and data across sites, and neither platform sees or dictates its programming beyond what it explicitly shares.

By contrast, a bowling-shoe agent is provided by the platform itself. These agents enjoy deep integration, privileged signals, and low setup costs, but are less portable, may steer outcomes toward the platform's interests, and can contribute to platform lock-in.

Orthogonal to ownership is specialization. We posit that there are likely to be two types of AI agents along this dimension: *horizontal* and *vertical*. Horizontal agents are generalists that span many tasks and surfaces, leveraging a single memory/preferences layer across markets and arbitraging options when beneficial, but they may lack domain-specific tooling or compliance features.

Vertical agents, on the other hand, specialize in a narrow domain (e.g., tax filing, job search, travel) or even a specific platform workflow, trading breadth for depth—stronger performance, richer integrations, and tailored guardrails—at the cost of portability and reuse.

Taken together, these two axes yield four archetypes, shown in Table 2, that structure the ensuing analysis.

Specialization / Ownership	Horizontal	Vertical
Bring-your-own agent	**Features:** User-controlled agent; not operated by the platform; carries cross-site memory/preferences; uses public APIs/standard interfaces; limited privileged hooks. **Pros:** Portable across markets; strong user alignment; privacy; can compare/arbitrate across platforms. **Cons:** Possible throttling/degraded access; weaker first-party data/tools; setup/subscription/compute cost.	**Features:** User-controlled specialist for a narrow domain (e.g., tax, jobs, travel); interoperates across competing platforms within that domain; third-party (not platform-run). **Pros:** Higher task performance than BYO-horizontal; retains user alignment; reusable across multiple platforms in-domain. **Cons:** Still lacks platform-privileged integrations; must track per-platform APIs/policies; limited reuse outside the domain.

Bowling-shoe agent	**Features:** Platform-operated generalist embedded in OS/app/site; convenient defaults; first-party telemetry and UI control. **Pros:** Low user friction; strong latency/reliability; access to proprietary features/tools. **Cons:** Limited portability; steering/self-preferencing risk; lock-in; weaker inspectability.	**Features:** Platform-operated specialist tightly integrated with domain tooling, policies, and datasets; optimized end-to-end flows with guardrails/compliance. **Pros:** Best performance on owning platform; richest domain functionality; full UX control for the platform. **Cons:** Highest degree of steering/lock-in; least transparent/inspectable; cross-platform substitution discouraged.

The choice between bring-your-own-agent and platform-provided agent models presents trade-offs from the consumer perspective. Bringing your own agent offers the advantages of perfect (or at least better) alignment with personal preferences and privacy, cross-platform functionality, and access to detailed personal information, allowing users to maintain consistent experiences across different services. However, this approach comes with significant costs and the risk of being outperformed by more sophisticated platform-specific alternatives. Eventually, platforms might not even allow direct access to their technologies without accessing their agent intermediaries (Rothschild et al., 2025). Conversely, platform-provided agents eliminate technical complexity and may be better if they are trained on platform specific data. But bowling-shoe agents may also sacrifice perfect alignment with individual needs, either due to explicit self-preferencing or due to simply not considering options offered on other platforms.

From a platform provider's perspective, the choice of AI agent model creates strategic trade-offs around control, costs, and competitive positioning. The bring-your-own-agent model reduces computational and hosting expenses for platforms while minimizing liability and simplifying API maintenance, but it sacrifices usage insights, opportunities for optimizing the user experience, and potentially profitable opportunities to steer consumers to the platform's preferred options. Furthermore, the bring-your-own-agent model reduces platform lock-in effects.

Platform-provided agents allow companies to maintain control over the user experience, benefit from their own R&D investments, and potentially steer users toward preferred options, but require substantial hosting and computation resources while potentially suppressing consumer demand due to alignment concerns. This fundamental tension between user autonomy and platform control mirrors broader debates in digital markets about the optimal balance between personalization, cost, and market power. We discuss these issues later in Section 6.2.

We can also imagine more complex structures, such as Anthropic's horizontal agent being integrated with an iPhone, or asking for help from Walmart's vertical agent,[6] or interacting directly

[6] Walmart has announced plans for an AI shopping agent (Bousquette, 2025).

with a seller's agent while dis-intermediating Walmart altogether. Agent to agent interfaces and agent-only storefronts may also proliferate. Interfaces that are made primarily for agents may be useful because of their speed, and the ability of the interface provider to tailor information to the agent. From a consumer's perspective, however, the agentic interaction becomes less inspectable.

4.1. Pricing of AI Agents

Another fascinating question in the supply argument is how agent services will be priced. Human agent services are often priced as a percentage of the transaction value. The sums involved can be high, especially in situations where the agent is paid to engage in an adversarial situation against another agent, and where one side wins and another loses. The high sums are due to two types of market forces: that the "best" agents are scarce (Rosen, 1981) and that high-powered incentives result in extraordinary effort (Holmstrom, 1979, Lazear and Rosen, 1981).

In contrast, supply of AI agents is virtually unconstrained, as software can be costlessly copied and the only marginal cost is the cost of maintenance, which is constantly declining. Whereas humans derive utility from pecuniary compensation, AI agents do not. They can be programmed to pursue monetary payoffs instrumentally, but they do not derive utility from money itself. Both factors point to diminished pricing power for AI vs human agents. One factor that may push in the opposite direction of high prices is if the quality of the agent's performance is a function of the compute used. In that case, we can imagine prices depending on the amount of compute given to the agent, which could be large for high-stakes transactions.

If there are strongly diminishing marginal returns to agent quality, then we can expect there to be competition in a manner that is similar to other digital services such as search and social media. In this scenario, the AI agent may be fully ad-supported. Providers may also make agents free, but monetize off complementary goods (e.g., phones) and services (e.g., delivery). It is also possible to see more sophisticated pricing schemes like versioning: a free but throttled agent for price-sensitive consumers and a full-featured Pro plan for less elastic users, with the free tier subsidizing off of the Pro plan.

In summary, there are many open questions about which types of agents will win out in equilibrium, and how they will be priced. In the next section, we consider how the widespread use of AI agents affects the economy.

5. Equilibrium Effects of AI Agency Under the Status Quo

To understand how AI agents will reshape markets, we can examine their effects within existing market structures before considering how market design itself might evolve. Consider a prototypical e-commerce market where AI agents enable users to consider all available options with complete information access. In such environments, sophisticated AI agents would likely prove resistant to nudges and advertising that lacks informational or signaling value, fundamentally altering competitive dynamics.

One of the most immediate impacts of AI agents will be their ability to dramatically lower search costs. Search-theoretic models have long emphasized how substantial cognitive and time-based costs prevent consumers from gathering and processing information necessary for optimal purchasing decisions. AI agents, with their superior computational capabilities, can rapidly collect, analyze, and compare extensive product data across markets, significantly mitigating these frictions and enhancing static allocative efficiency.

These static improvements would generate important dynamic effects over time. As AI agents consistently direct demand toward higher-quality or better-value offerings, producers would receive clearer market signals about consumer preferences. This feedback mechanism would incentivize firms to invest in producing what consumers want rather than what they can be persuaded to buy through marketing or by exploiting cognitive limitations. It would also incentivize firms to allow for more customization in their products, given that the costs of customization and discovery will drop. The resulting shift in production patterns would compound initial allocative gains, creating a virtuous cycle of better-targeted supply meeting more accurately expressed demand.

Advanced AI agents will also enable new pricing mechanisms through enhanced preference elicitation and revelation.[7] When consumers willingly and accurately reveal their preferences to trusted AI agents, firms gain access to richer, more detailed preference information. This enhanced informational access enables personalized pricing strategies that can significantly improve price discovery and market efficiency. While personalized pricing can benefit consumers by closely matching products to individual preferences, it may also lead to inequitable outcomes if firms exploit consumer-specific price elasticity to maximize profits at the expense of consumer welfare. In fact, agents may strategically hold back some information about a customer in order to obtain more favorable prices. Regulation will likely be necessary to facilitate and maintain consumer protections in these cases. We discuss in detail different regulatory avenues in the last section of this chapter.

AI agents will affect the prevalence and dynamics of bargaining because their opportunity cost is not human time. For the principal, delay is costly—time and attention are scarce resources, so classic bargaining models assume impatience and a preference to conclude negotiations quickly. For the agent, the binding cost is compute and API usage rather than time, and those costs are trending down. As a result, even if the principal prefers timely resolution, agents can initiate negotiations earlier and persist longer (e.g., begin negotiating summer-2027 rentals in January 2026).

It is an open question whether the use of AI agents will cause lower price dispersion or reduce other economic rents. Many economists initially predicted that the Internet would eliminate price dispersion in markets, reasoning that decreased search costs would create conditions similar to Bertrand competition. However, this prediction largely failed to materialize. A sizable theoretical and empirical literature in economics has investigated the reasons for continual price dispersion. When products are differentiated, lower search costs can paradoxically lead to higher prices and greater price dispersion (Ellison and Ellison, 2018, Kaye, 2025) as matching improves along the taste dimension. Furthermore, the internet has enabled a variety of obfuscation tactics that have counteracted the expected benefits of lower search costs, preserving or even increasing equilibrium price dispersion (Ellison and Ellison, 2009). Shopping by AI agents may be subject to similar forces.

[7] AI agents may also act as pricing agents producing similar complications as in Calvano et al. (2020) and Brown and MacKay (2023).

Perhaps the best case for AI agents reducing economic rents is in markets where firms exploit behavioral biases and bounded rationality. For example, many consumers choose suboptimal contracts given their expected usage patterns, leading to rents for firms (e.g, Grubb, 2009 for cell phone contracts). Since AI agents can make more rational decisions, such rent-extraction strategies would become far less viable, pushing markets closer to the competitive ideal.

However, mass proliferation of agents does not guarantee efficient market operation, especially in cases characterized by congestion, imperfectly aligned agents, and asymmetric information. For example, reducing the cost of applying to a job with a customized cover letter will result in many more job applications for each position, making it harder for employers to pick the right employee (Wiles and Horton, 2025). This highlights the critical role that market design will play in ensuring positive-sum outcomes as AI agents become ubiquitous in economic transactions.

6. Market Design for AI Agents

As AI agents become ubiquitous in economic transactions, markets themselves will need fundamental reconfiguration to take advantage of their capabilities and accommodate the unique challenges they offer. This transformation extends beyond policy adjustments to encompass technical infrastructure, identity verification systems, and entirely new market mechanisms that use agents' superior computational abilities. In this section, we discuss aspects of market design, as well as their application to several important industries.

When economists consider market reconfiguration, they typically focus on policy changes such as pricing or information structures. However, the advent of AI agents also requires more technical considerations of whether digital platforms can effectively accommodate agent interactions. Take for example the problem of website traffic. Currently, users do not pay websites per http request. This is due to a combination of the low costs to the domain of an individual request, and due to technical challenges. However, as agentic browsing becomes ubiquitous, the overall volume of traffic is likely to explode, with website owners facing the costs. Ultimately, the new traffic and cost profile may call for agent-first surfaces—authenticated, rate-limited APIs with machine-readable pricing and consent signals—rather than human-oriented pages.

We expect new conventions and capabilities to emerge, similar to the existing *robots.txt* standard, where sites specify how agents should interact with their platforms and what activities are permitted or prohibited. For example, Cloudflare is a content delivery network provider which helps domains manage traffic. It recently introduced 'pay-per-crawl', which is a feature that allows website owners to charge agents when crawling their website (Allen and Newton, 2025). This new capability for websites follows the Pigouvian intuition that market participants should pay for the externalities they impose on others. Additional technical evolution may manifest itself in the distinction between agent-oriented and person-oriented interfaces. Websites will develop parallel access methods: streamlined, data-rich interfaces optimized for agent consumption alongside traditional user interfaces designed for human interaction.

6.1. Identity

A fundamental challenge of implementing market designs emerges from the reality that most internet activity will eventually originate from agents rather than humans. These AI agents will be able to effectively mimic humans in many situations. For example, ads are priced based on impressions and clicks, but AI agents will also be able to 'see' ads and click on them. Similarly, social media companies and their users would like to clearly verify that content is produced by a specific human rather than a bot. Sybil attacks—where single entities create multiple fake identities to gain disproportionate network control—become more difficult to detect in agent-mediated environments. This necessitates new approaches to digital accountability, potentially including cryptographic proofs, digital IDs, or decentralized verification platforms to ensure system safety and reliability.

Of particular importance is identity verification, both for humans and for AI systems that are acting on behalf of specific humans. Two broad classes of solutions emerge: walled-garden approaches where platform gatekeepers exclude potentially malicious actors, and open systems with robust verification mechanisms. In the walled-garden approach, platforms require a log-in prior to interacting with content. This approach is imperfect. A person can launch an agent after logging-in, creating spam and malicious content. If the platform bans this person, this same person can create additional accounts.

Another approach is a proof-of-personhood system, which tries to create a network of unique humans. For example, the World Foundation has created a biometric technology that uses the human iris to uniquely identify humans and to give them access to an app that can be used to prove this uniqueness (World Foundation, 2023). Proof of humanhood solutions can prevent sybil attacks and can preserve privacy, but require a larger restructuring of existing systems and mass adoption to be truly transformative.

Identity systems will also need to be combined with verified credentials and reputation mechanisms. For example, consider an advertiser wishing to target humans with particular demographics. Some of these demographics are verifiable using a digital government identity, and this identity can be used to cryptographically prove demographics to a particular advertiser. Reputation mechanisms can also be designed to be attached to particular identities in a similar manner. Combined, identity, credentials, and reputation can enable more sophisticated market interactions, including targeted deals and negotiations over small purchases that would previously have been economically infeasible.

Absent an identity requirement, the anonymity enabled by AI agents creates both opportunities and challenges for market design. Agents can maintain negotiating anonymity, potentially helping users achieve fairer outcomes by avoiding premature information revelation. The classic example of Disney's anonymous land purchases for Walt Disney World demonstrates how strategic anonymity can prevent price manipulation. However, market designers must carefully balance anonymity benefits against verification needs, considering whether multiple personas linked to single individuals should be permitted and how to prevent harmful duplicity (Buterin, 2025).

6.2. Changes in Existing Platform Design

Platforms of all types, including those for e-commerce, job search, and social media will need to change in response to AI agents. We discuss two types of changes, that to the content served by the platform and that to the content created by humans and AI agents. Of course, both the content consumption and content creation sides of the market are related in equilibrium.

First, consider the fact that AI agents may consume content such as social media posts and search results in e-commerce. AI will serve as a sophisticated filter between humans and the broader digital environment. Much of today's online content combines user-desired material with content designed primarily to capture attention or generate revenue, from native advertising to engagement-optimized recommendation algorithms. AI agents could review incoming information streams and selectively transmit content based on user preferences and actual utility, substantially reducing exposure to irrelevant material.

This filtering function poses significant challenges to existing digital business models that rely on bundling content with advertising or engagement-driven recommendations. For example, much of digital design uses choice architecture to influence users into making the actions preferred by the designer. Rational AI agents browsing on behalf of users are unlikely to be influenced by these nudges. Similarly, advertising is often characterized by puffery, and selective information disclosure. AI agents may be much less influenced by these types of ads than humans.

As agents proliferate, platforms will need to adjust. For example, platforms will try to design information architectures that specifically influence agents, something that is already evident with the nascent field of optimization of content for LLM consumption. Over time, we may see a shift to alternative monetization models, such as subscription services, which have less dependence on advertising.

Second, consider the content creation ability of AI agents. AI agents will be able to create and like social media posts, create artistic works, send direct messages, offer to buy items, and apply to jobs. Since the costs of these actions is low, it may be individually rational for humans to ask their AI agents to do these actions en masse (Goldberg and Lam, 2025). Yet part of the value of these actions is that they serve as signals.

In response to content proliferation, platforms will be forced to adapt rules restoring costly signaling and credible verification that content was created by a human. For example, posters on social media may be required to pay per post, so that the platform does not get flooded with low-quality or undifferentiated content. In e-commerce, fraud is already a first-order concern due to fake listings and credit card chargebacks. We expect platforms to add monetary costs or other limits on these activities.

Merely the fact that a piece of content was produced by a human vs an AI may be important for signaling and consumption utility (Longoni et al., 2022, Rae, 2024). We've already discussed the potential applications of digital identity in a world of AI agents. Platforms will face a choice of which identity, if any, to require, and in how to use identity related information.

6.3. Enabling Previously Impractical Market Designs

The near-zero marginal cost of agent operation makes previously impractical, yet appealing, market mechanisms economically viable. Economic theory has long identified superior market designs that remain underutilized due to implementation costs, particularly those requiring fine-grained preference information impossible or inefficient for humans to provide at scale.

Consider the deferred acceptance algorithm, which offers stability and strategy-proofness for proposing parties (Gale and Shapley, 1962). While many labor markets and matching services could benefit from such mechanisms, they remain largely underutilized, likely because they require comprehensive preference rankings that are cognitively difficult or economically costly to provide. Instead, matching platforms (e.g., Upwork, Tinder, etc.) often rely on far simpler recommendation systems—providing all participants with the same "best" options—leaving many with participants with unstable or entirely unmatched outcomes. AI agents can help solve such inefficiencies.

Recent empirical work demonstrates the potential: when crowdsourcing workers provide natural language descriptions of their task preferences, LLM-generated rankings over large choice sets prove more accurate than human-generated rankings over smaller sets (Rusak et al., forthcoming). This suggests that cognitive limitations and time costs rather than preference complexity represent the binding constraint.

With AI-derived preferences, markets could implement sophisticated matching algorithms requiring preferences over thousands of alternatives. Labor markets could have both job seekers and employers provide natural language descriptions of their preferences, enabling comprehensive rankings that support equilibrium matching algorithms superior to traditional recommendation systems. This capability extends beyond labor markets to any matching context where the complexity of optimal allocation exceeds human computational capacity.

AI agents also enable market designs to maintain privacy in ways that would otherwise be very difficult with humans. Like when principals face strategic concerns when seeking sensitive information, as inquiries may inadvertently send negative signals. For example, a job seeker may hesitate to directly inquire about maternity leave policies to avoid negative signaling to employers, despite having a legitimate interest in obtaining this information.

By delegating sensitive questions to AI agents, parties can credibly commit to privacy-preserving interactions. Both sides can precommit their agents to pose all legally permissible questions without penalty, with agents disclosing only relevant responses. These mechanisms may effectively separate sensitive inquiries from signaling concerns, enhancing transparency and efficiency across sensitive contexts.

7. Regulation

As AI agents become increasingly integrated into critical decision-making processes, regulatory frameworks will need to evolve to address emerging challenges around market power, autonomy, liability, privacy, and intellectual property rights. Current policy efforts range from targeted municipal ordinances to comprehensive federal initiatives, while technical safeguards continue to develop alongside these regulatory approaches. The rapid pace of AI advancement often outstrips the ability of existing regulations to provide adequate protection, creating a need for adaptive regulatory frameworks that can keep pace with technological development.

Market Power: A central concern in AI regulation is the concentration of market power among a handful of large firms. These companies control the vast computational resources, proprietary datasets, and distribution channels necessary to develop frontier models, creating high barriers to entry for competitors. Without regulatory measures, such dominance risks entrenching incumbents, limiting innovation, and reducing consumer choice. Effective regulation must therefore address not only safety and accountability but also competitive dynamics—for example, by ensuring access to essential infrastructure, promoting interoperability standards, and preventing exclusionary practices. At the same time, regulators must guard against regulatory capture, where dominant firms shape rules in ways that reinforce their advantage under the guise of compliance.

Current developments illustrate the stakes. As of September 2025, antitrust scrutiny of large AI firms has intensified, including investigations into Google's market practices (U.S. Department of Justice, 2025). These cases highlight the delicate balance regulators face: while underregulation risks leaving unchecked monopolistic behavior, overregulation could inadvertently stifle technological innovation and slow beneficial applications of AI.

Autonomy and Liability: With AI agents acting on behalf of humans, the question of who bears responsibility when things go wrong becomes unavoidable. Whether accountability should rest with the human who delegates the final decision to an AI or the human who acts on another party's AI-generated output (i.e., under negligence liability), or with the firms that develop or deploy such systems regardless of fault (i.e., strict liability), remains contested. This choice raises fundamental concerns about how to allocate the costs of harm in a way that protects users without choking off innovation.

The European Union's (EU) recent product liability directive marks the most advanced regulatory attempt to grapple with these challenges to date, explicitly extending liability to digital goods, software, and ongoing updates (European Union, 2024). By recognizing that the traditional boundary between a "finished product" and its subsequent behavior no longer holds in the age of AI—as these systems learn, adapt, and sometimes act in ways that surprise even their creators—the directive exemplifies how adjustments to existing tort frameworks, rather than the creation of entirely new ones, may become the central lever shaping the trajectory of AI development and adoption.

Even with clearer ex-post liability measures, AI systems remain vulnerable to adversarial manipulation. Jailbreaking attacks allow users to circumvent safety measures and access system features in unintended ways, representing a persistent challenge for current chatbot technology (Shen et al., 2024). The vulnerabilities become particularly concerning in high-stakes applications like hiring, where malicious actors might exploit AI interviewers through carefully crafted prompts to gain unfair advantages.

Privacy: These failure modes magnify privacy risks: once guardrails are bypassed, agents can pull data across contexts, retain it beyond its intended use, or infer it from seemingly benign traces. Training and adaptation on sensitive user data—often without explicit consent—heightens these risks and raises basic questions about data ownership and control, especially in employment, housing, and other high-stakes domains.

A particular privacy risk stems from inadvertent training on private or sensitive data. For example, a model trained based on Bob's social media posts may memorize information about Bob. When Amy writes a prompt relating to Bob, information about Bob may surface in unpredictable and potentially damaging ways. In addition to the leakage of private information memorized in training, another risk may simply be sophisticated inference from usage traces like late-night browsing or vitamin purchases. This information can be repurposed in other contexts (e.g., lower salary offers based on perceived pregnancy or higher insurance premiums based on inferred health risks). While there are already various efforts to regulate data use, such as the General Data Protection Regulation (2016) in the European Union, these may need to be adapted in response to generative AI and automated decision-making (Workforce Bulletin, 2023).

Intellectual Property: Data protection concerns extend even beyond privacy and into intellectual property. Generative AI models are trained on vast repositories of creative content, much of it scraped without authorization. Once such data has been incorporated, "unlearning" is technically difficult, raising questions about compensation for artists, writers, and other creators.

Platforms are already taking actions to block AI agents from accessing their content via litigation (The New York Times Co. v. Microsoft Corp., 2023, Smith, 2025), robots rules, and tighter terms of service. Yet more capable agents can emulate human browsing or operate on the principal's device to conduct certain actions. This creates a real tension. While users want tools that act on their behalf across the web, platforms want enforceable control and intellectual property protection.

Rather than blocking alone, policy can set shared rails: (i) universal opt-out registries that agents must check; (ii) pooled/collective licensing that compensates creators; (iii) provenance and authenticity standards with tamper-evident metadata; and (iv) periodic audits and penalties for non-compliance. Together, these as well as other tools align training and use with both creator rights and user autonomy.

8. Conclusion

The capacity of AI agents to dramatically reduce transaction costs through automated intermediation could unlock new forms of market participation, enable previously infeasible mechanisms, and push allocative efficiency closer to theoretical ideals. Yet the same forces that make agents attractive—their tireless persistence, computational superiority, and negligible marginal costs—also threaten to overwhelm existing market structures. The ultimate impact will depend critically on collective choices regarding agent design, market structures, and regulatory frameworks. Economists are uniquely positioned to offer both theoretical insights and empirical analysis to help to guide this transition toward welfare-enhancing rather than merely rent-redistributive outcomes.

Acknowledgments

We thank David Rothschild for his detailed comments.

References

Allen, Will, and Simon Newton. “Introducing Pay Per Crawl: Enabling Content Owners to Charge AI Crawlers for Access.” *Cloudflare Blog*, July 1, 2025. https://blog.cloudflare.com/introducing-pay-per-crawl/.

Allouah, Amine, Omar Besbes, Josué D. Figueroa, Yash Kanoria, and Akshit Kumar. 2025. “What Is Your AI Agent Buying? Evaluation, Implications and Emerging Questions for Agentic E-Commerce.” arXiv preprint arXiv:2508.02630.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. “Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.” American Economic Review 94 (4): 991–1013. https://doi.org/10.1257/0002828042002561.

Bostrom, Nick. 2014. Superintelligence: Paths, Dangers, Strategies.

Bousquette, Isabelle. 2025. “Walmart Is Preparing to Welcome Its Next Customer: The AI Shopping Agent.” *The Wall Street Journal*, May 2025.

Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. “General Purpose Technologies ‘Engines of Growth’?” Journal of Econometrics 65 (1): 83–108.

Brown, Zach Y., and Alexander MacKay. 2023. “Competition in Pricing Algorithms.” American Economic Journal: Microeconomics 15 (2): 109–56.

Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond. 2025. “Generative AI at Work.” The Quarterly Journal of Economics 140 (2): 889–942.

Buterin, Vitalik. 2025. “Does digital ID have risks even if it's ZK-wrapped?” https://vitalik.eth.limo/general/2025/06/28/zkid.html

California Consumer Privacy Act of 2018. *Cal. Civ. Code § 1798.100 et seq., as amended*. “California Consumer Privacy Act (CCPA).” California Attorney General. Updated March 13, 2024. Accessed September 15, 2025. https://oag.ca.gov/privacy/ccpa.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolo, and Sergio Pastorello. 2020. “Artificial Intelligence, Algorithmic Pricing, and Collusion.” American Economic Review 110 (10): 3267–97.

Cesaratto, Brian G., Frances M. Green, Nathaniel M. Glasser, and América Garza. 2025. “California’s AI Revolution: Proposed CPPA Regulations Target Automated Decision Making.” Workforce Bulletin.

Chatterji, Aaron, Thomas Cunningham, David J. Deming, Zoe Hitzig, Christopher Ong, Carl Yan Shan, and Kevin Wadman. *How People Use ChatGPT*. NBER Working Paper 34255. Cambridge, MA: National Bureau of Economic Research, 2025. http://www.nber.org/papers/w34255.

Christian, Brian. 2020. The Alignment Problem: Machine Learning and Human Values. New York: W. W. Norton & Company.

Citron, Dave. 2024. “Try Deep Research and Our New Experimental Model in Gemini, Your AI Assistant.” *Google Blog*, December 11, 2024. https://blog.google/products/gemini/google-gemini-deep-research/

Coase, R. H. 1937. “The Nature of the Firm.” Economica 4: 386–405. https://doi.org/10.1111/j.1468-0335.1937.tb00002.x.

Cui, Zheyuan Kevin, Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz. 2025. “The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers.” SSRN Working Paper 4945566.

Dammu, Preetam Prabhu Srikar, Omar Alonso, and Barbara Poblete. 2025. “A Shopping Agent for Addressing Subjective Product Needs.” In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining (WSDM ’25)*.

Dobbala, M. K., and M. S. S. Lingolu. 2024. “Conversational AI and Chatbots: Enhancing User Experience on Websites.” American Journal of Computer Science and Technology 7 (3): 11.

Ellison, Glenn, and Sara Fisher Ellison. 2009. “Search, Obfuscation, and Price Elasticities on the Internet.” Econometrica 77 (2): 427–52.

Ellison, Glenn, and Sara Fisher Ellison. 2018. Match Quality, Search, and the Internet Market for Used Books. NBER Working Paper w24197. Cambridge, MA: National Bureau of Economic Research.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. “GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models.” arXiv preprint arXiv:2303.10130.

European Union. *Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC*. OJ L 2024/2853, 18 November 2024. http://data.europa.eu/eli/dir/2024/2853/oj

Fradkin, Andrey. 2025. “Demand for LLMs: Descriptive Evidence on Substitution, Market Expansion, and Multihoming.” arXiv preprint arXiv:2504.15440.

Gaebler, Johann D., Sharad Goel, Aziz Huq, and Prasanna Tambe. 2024. “Auditing Large Language Models for Race & Gender Disparities: Implications for Artificial Intelligence–Based Hiring.” Behavioral Science & Policy 10 (2): 46–55.

Gale, David, and Lloyd S. Shapley. 1962. “College Admissions and the Stability of Marriage.” The American Mathematical Monthly 69 (1): 9–15. https://doi.org/10.2307/2312726.

General Data Protection Regulation (GDPR). 2016. *Eur. Parl. & Council Regulation 2016/679, OJ L 119*, 4 May 2016; corrected in OJ L 127, 23 May 2018. Accessed September 15, 2025. https://gdpr-info.eu

Goldberg, Samuel, and H. Tai Lam. 2025. “Generative AI in Equilibrium: Evidence from a Creative Goods Marketplace.” Equilibrium: Evidence from a Creative Goods Marketplace (February 24, 2025).

Grubb, Michael D. 2009. “Selling to Overconfident Consumers.” American Economic Review 99 (5): 1770–807.

Hadfield, Gillian K., and Andrew Koh. 2025. “An Economy of AI Agents.” *arXiv*, September 1. https://arxiv.org/abs/2509.01063. doi:10.48550/arXiv.2509.01063.

Heitz, Miriam, Stefan König, and Torsten Eymann. 2010. “Reputation in Multi-Agent Systems and the Incentives to Provide Feedback.” In German Conference on Multiagent System Technologies. Berlin, Heidelberg: Springer.

Hilel, A., Shenfeld, I., Andreas, J., Choshen, L. 2025. “LLM Hypnosis: Exploiting User Feedback for Unauthorized Knowledge Injection to All Users.” arXiv preprint arXiv:2507.02850.

Holmström, Bengt. 1979. “Moral Hazard and Observability.” The Bell Journal of Economics 10 (1): 74–91. https://doi.org/10.2307/3003320.

Immorlica, Nicole, Brendan Lucier, and Aleksandrs Slivkins. 2024. “Generative AI as Economic Agents.” *arXiv* preprint arXiv:2406.00477.

Jabarian, B., & Henkel, L. (2025). *Voice AI in firms: A natural field experiment on automated job interviews* (Job market paper). University of Chicago & Erasmus University Rotterdam.

Ji, Ziwei, Tiezheng Yu, Yan Xu, Nayeon Lee, Etsuko Ishii, and Pascale Fung. 2023. “Towards Mitigating LLM Hallucination via Self-Reflection.” Findings of the Association for Computational Linguistics: EMNLP 2023.

Kalai, Adam Tauman, Ofir Nachum, Santosh S. Vempala, and Edwin Zhang. 2025. “Why Language Models Hallucinate.” arXiv preprint arXiv:2509.04664.

Kaye, Aaron. 2024. The Personalization Paradox: Welfare Effects of Personalized Recommendations in Two-Sided Digital Markets. Working Paper.

Kessler, Sarah. “Employers Are Buried in A.I.-Generated Résumés.” *New York Times*, June 21, 2025. Updated June 23, 2025. https://www.nytimes.com/2025/06/21/business/dealbook/ai-job-applications.html

Lazear, Edward P., and Sherwin Rosen. 1981. “Rank-Order Tournaments as Optimum Labor Contracts.” Journal of Political Economy 89 (5): 841–64.

Liang, Annie. "Artificial Intelligence Clones." *arXiv preprint arXiv:2501.16996* (2025).

Longoni, C., Fradkin, A., Cian, L., & Pennycook, G. (2022, June). News from generative artificial intelligence is believed less. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 97-106).

Manakul, Potsawee, Adian Liusie, and Mark J. F. Gales. 2023. “SelfCheckGPT: Zero-Resource Black-Box Hallucination Detection for Generative Large Language Models.” arXiv preprint arXiv:2303.08896.

New York City Council. 2023. “Local Law 144 of 2021: Automated Employment Decision Tools.” NYC Administrative Code, Title 20, Chapter 5, Subchapter 25.

The New York Times Co. v. Microsoft Corp. 2023. No. 1:23-cv-11195 (S.D.N.Y., December 27). Complaint. PDF available at https://nytco-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf. Accessed September 9, 2025.

Ouyang, Long, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang et al. 2022. “Training Language Models to Follow Instructions with Human Feedback.” Advances in Neural Information Processing Systems 35: 27730–44.

Rae, Irene. 2024. “The Effects of Perceived AI Use on Content Perceptions.” In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems.

Rosen, Sherwin. 1981. “The Economics of Superstars.” The American Economic Review 71 (5): 845–58.

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. 2025. “The Agentic Economy.” *arXiv* preprint arXiv:2505.15799.

Rusak, Gili, Benjamin S. Manning, and John J. Horton. 2025. “AI Agents Can Enable Superior Market Designs.” https://conference.iza.org/DATA_2025/manning_b36140.pdf.

Shen, Xinyue, Zeyuan Chen, Michael Backes, Yun Shen, and Yang Zhang. 2024. “‘Do Anything Now’: Characterizing and Evaluating In-the-Wild Jailbreak Prompts on Large Language Models.” In Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security.

Smith, Allison. “Shopify Quietly Sets Boundaries for AI Agents on Merchant Sites.” *Modern Retail*, July 14, 2025. https://www.modernretail.co/technology/shopify-has-quietly-set-boundaries-for-buy-for-me-ai-bots-on-merchant-sites/.

Spatharioti, Sofia Eleni, David Rothschild, Daniel G. Goldstein, and Jake M. Hofman. 2025. “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*, 1–15.

U.S. Department of Justice, Office of Public Affairs. “Department of Justice Prevails in Landmark Antitrust Case Against Google,” April 17, 2025. https://www.justice.gov/opa/pr/department-justice-prevails-landmark-antitrust-case-against-google

Wiles, Emma, Zanele Munyikwa, and John Horton. 2025. “Algorithmic Writing Assistance on Jobseekers’ Resumes Increases Hires.” Management Science.

Wiles, Emma, and John J. Horton. "Generative ai and labor market matching efficiency." *Available at SSRN 5187344*(2025).

World Foundation. 2023. “Iris Recognition Inference System.” World.org, June 5, 2023. https://world.org/blog/engineering/iris-recognition-inference-system. Accessed September 14, 2025.

Wu, Junhui, Daniel Balliet, and Paul A. M. Van Lange. 2016. “Reputation, Gossip, and Human Cooperation.” Social and Personality Psychology Compass 10 (6): 350–64.

Zeff, Maxwell. 2025. “Google Rolls Out Project Mariner, Its Web-Browsing AI Agent.” *TechCrunch*, May 2025.

Zhang, Biao, Zhongtao Liu, Colin Cherry, and Orhan Firat. 2024. “When Scaling Meets LLM Finetuning: The Effect of Data, Model and Finetuning Method.” arXiv preprint arXiv:2402.17193.

Zhu, Shenzhe, Jiao Sun, Yi Nian, Tobin South, Alex Pentland, and Jiaxin Pei. 2025. “The Automated but Risky Game: Modeling Agent-to-Agent Negotiations and Transactions in Consumer Markets.” arXiv preprint arXiv:2506.00073.