Strategy & Competitiveness

AI Transformation: Economic Frameworks
for the Intelligence Age

Michael Ghilissen
michael.ghilissen@uliege.be

Executive Summary

We stand at a singular inflection point. Artificial intelligence is not simply another productivity-enhancing technology layered onto existing economic structures. It is, in the words of the leading researchers whose work this paper synthesizes, a potential source of transformative change — one that compresses centuries of innovation into decades, shifts the locus of decision-making from humans to machines, and strains every economic institution we have built since the Industrial Revolution.

This paper distills the emerging economic scholarship on AI transformation into a coherent set of frameworks for business leaders and policymakers. Drawing on eight years of research from the National Bureau of Economic Research's Economics of Artificial Intelligence Initiative — and especially the landmark 2025 volume A Research Agenda for the Economics of Transformative AI — it addresses five interconnected challenges:

- How AI reshapes the fundamental economics of innovation, measurement, and growth
- How AI disrupts markets, competition, and the boundaries of the firm
- How AI transforms labor markets, income distribution, and human welfare
- How AI changes knowledge creation, information flows, and systemic risk
- What policy frameworks are adequate to govern AI's fiscal and existential dimensions

The central thesis is this: AI's economic impact is not merely quantitative — more output, faster processes — but qualitative. It changes what can be measured, who bears economic risk, how knowledge is created, and ultimately what work means for human beings. Business leaders who see AI only through the lens of efficiency gains will be systematically surprised by the transformations ahead.

A Word on Sources

This paper synthesizes the contributions of Agrawal, Gans, Goldfarb, Brynjolfsson, Korinek, Ben Jones, Chad Jones, Manning, Restrepo, Stevenson, Athey, Stiglitz, Mullainathan, Hadfield, and many colleagues whose collective output constitutes the most rigorous economic treatment of AI to date. All specific attribution is noted in the text. Readers are strongly encouraged to consult the primary sources for technical depth.

Introduction: A New Kind of General Purpose Technology

In 2017, Daniel Kahneman offered a provocation that has guided the field of Economics of AI ever since. The most important economic question about AI, he suggested, was not how much value it would create, but how it would change the process of prediction — and therefore the entire structure of decision-making under uncertainty. Eight years later, that insight looks prescient, even beyond Kahneman's framing.

AI is, in the taxonomy of economic historians, a General Purpose Technology: a foundational innovation whose value lies not in what it does directly, but in how it enables everything else. Electricity, the steam engine, the internet — each reshaped production, organization, and ultimately society over decades, not years. What is distinctive about AI is the speed, breadth, and nature of the disruption it portends.

> *"The question is not whether AI will transform the economy, but whether our economic institutions — our measurement systems, tax bases, labor markets, and governance structures — are adequate to manage that transformation."*
> **— Brynjolfsson, Korinek & Agrawal, 2025**

Transformative AI — meaning AI systems capable not merely of augmenting human tasks but of autonomously performing cognitive work across a wide range of domains — introduces at least three qualitative breaks from prior technological waves:

- It is a technology that improves itself. Unlike the steam engine, AI systems can, under certain conditions, generate better AI systems, creating the possibility of recursive self-improvement that standard economic models cannot capture.
- It directly substitutes for cognitive labor. Previous automation displaced physical effort; AI displaces thinking, judgment, and creativity — precisely the capacities that gave knowledge workers their economic premium.
- It is simultaneously infrastructure, product, and actor. AI is not merely a tool that businesses use; through agentic systems, it is becoming a participant in economic transactions, blurring the line between firm, contract, and employee.

This paper proceeds in five parts, each corresponding to a major domain of economic inquiry. The goal is not encyclopedic coverage but the identification of the core frameworks that MBA-level leaders and strategists need to navigate the coming decade.

PART I

Foundations of Transformative AI Economics

Understanding AI's economic significance requires updating foundational concepts: what innovation means in an AI-enabled world, what growth theory predicts about AI's long-run impact, and — critically — whether our measurement systems are adequate to track what is actually happening.

1.1 Prediction, Uncertainty, and the New Economics of Decision

The most durable conceptual contribution of the NBER program comes from Agrawal, Gans, and Goldfarb's foundational framework, developed in their 2019 volume and refined through 2025. Their core insight is deceptively simple: AI dramatically reduces the cost of prediction.

In economic decision theory, every choice under uncertainty involves a prediction about future demand, competitor behavior, input costs, employee performance, and customer churn. Historically, prediction was expensive: it required human expertise, data collection, and statistical analysis. AI industrializes prediction, driving its cost toward zero.

This matters enormously because when the cost of an input falls precipitously, two things happen. First, we use much more of it — prediction becomes embedded in processes that were previously absent. Second, the value of complements to that input rises. As Agrawal, Gans, and Goldfarb argue in their 2025 update on what they call 'genius on demand,' when prediction becomes cheap, the scarce and valuable inputs shift to judgment — the human capacity to specify objectives, weigh incommensurable values, and decide what matters when data alone cannot.

Framework: The Prediction-Judgment Frontier

Think of every professional task as lying on a spectrum. At one extreme: pure prediction (classifying transactions, translating text, diagnosing images). At the other: pure judgment (setting strategy, resolving ethical dilemmas, managing creative ambiguity). AI shifts the frontier: tasks once requiring both prediction and judgment can now be decomposed. The prediction component is automated; the judgment component is elevated in importance. The strategic question for every organization is: where does our value creation sit on this frontier, and how is AI moving it?

The practical implication is a disaggregation of professional work that is only beginning. Diagnosis in medicine, legal research, financial analysis, and architectural design — each involves a mixture of prediction (pattern-matching from data) and judgment (defining goals, managing relationships,

accepting accountability). AI is unbundling these, which in turn unbundles traditional professions and firm structures.

1.2 Innovation at the Frontier: Ben Jones on the Bottleneck

One of the most consequential economic questions about AI concerns the innovation process itself. Ben Jones has spent a career documenting the rising burden of knowledge — the observation that reaching the frontier of any scientific or technological field requires ever-greater investment in education and specialization, pushing up the cost and time required for each incremental advance.

AI, Jones argues in his 2025 contribution, has the potential to fundamentally alter this dynamic. If AI can compress the apprenticeship phase — allowing researchers to absorb domain knowledge faster, generate hypotheses more rapidly, and test them at lower cost — the effective productivity of human researchers could rise dramatically. His bottleneck framework asks: where in the innovation pipeline is effort currently concentrated, and which bottlenecks can AI actually relieve?

Jones identifies three distinct modes of AI impact on innovation: AI as a tool that augments human researchers (search, synthesis, simulation); AI as a partial substitute that handles specific cognitive subtasks (literature review, data analysis, protocol generation); and AI as an autonomous research agent capable of formulating and testing hypotheses independently. These modes have very different economic implications — the third, in particular, could generate what Jones calls super linear growth: innovations beget innovations faster than researchers are added, breaking the historical pattern of diminishing returns to research investment.

For business leaders, the Jones framework translates to a concrete strategic question: how much of your organization's innovation pipeline is bottlenecked by prediction and synthesis tasks that AI can now perform, versus the genuinely creative, relationship-dependent, or tacit-knowledge tasks that remain human? Companies that correctly map this frontier will allocate R&D resources very differently from those that do not.

1.3 The Measurement Crisis: Coyle and Poquiz

There is a paradox at the heart of AI's economic history to date: a technology of extraordinary demonstrated capability has coincided with a period of disappointing measured productivity growth. Economists call this the AI productivity paradox, echoing the computer productivity paradox that Robert Solow identified in the 1980s.

Coyle and Poquiz's 2025 contribution argues that this paradox is partly real and partly an artifact of measurement failure. Our national accounting frameworks — GDP, TFP, sectoral productivity indices — were built to measure the production and consumption of physical goods and clearly priced services. They systematically undercount the value created when AI improves processes without changing prices, when it generates non-market goods (free services, improved quality), or when its benefits accrue as reduced costs and uncertainty rather than increased output.

> *"Our statistical frameworks cannot adequately measure what they cannot see. And they cannot see most of what AI is doing."*
> **— Coyle and Poquiz, 2025**

The measurement gap has direct strategic implications. Firms that invest heavily in AI may show depressed accounting returns — costs are visible, benefits are diffuse — while actually building substantial competitive advantage. Conversely, laggards who have not invested will not show a measurable productivity deficit until the gap becomes catastrophic. Standard financial metrics may actively mislead boards and investors about where value is being created and destroyed.

Korinek and Lockwood's 2025 analysis adds a fiscal dimension: traditional tax bases — corporate income, payroll, value-added — are built on measurable economic transactions. As AI shifts value creation toward intangibles, non-market goods, and automated processes that do not generate human incomes, the taxable base may shrink even as economic activity and well-being expand. This is not merely a fiscal curiosity; it has direct implications for the social contract that funds education, healthcare, and the safety net.

PART II

Markets, Competition, and Organization

AI does not operate in a vacuum; it operates within markets, firms, and competitive structures. Understanding how those structures change — and what new structures emerge — is among the most practically urgent areas of AI economics.

2.1 AI's Unique Economic Characteristics: Athey and Morton

Susan Athey and Fiona Morton's 2025 analysis of AI's competitive economics begins with a deceptively simple observation: AI has properties that violate most of the assumptions underlying standard theories of competition.

Consider the classic conditions for competitive markets: many buyers and sellers, free entry and exit, homogeneous products, and perfect information. AI markets tend toward the opposite: massive fixed costs and near-zero marginal costs create natural-monopoly dynamics; network effects and data feedback loops create self-reinforcing advantages for incumbents; the opacity of model training creates information asymmetries that are structurally distinct from those in traditional goods markets.

AI Characteristic	Economic Implication	Strategic/Policy Challenge
Non-rival consumption	The same AI model can serve millions of users simultaneously without degradation — unlike physical capital	AI costs are front-loaded in training; deployment is nearly free
Data network effects	More users generate more data, which improves the model, which attracts more users	Creates winner-take-most dynamics unlike most product markets
Feedback loops	AI recommendations shape behavior, which generates new training data that reinforces initial biases	Raises novel concerns about market manipulation and lock-in
Opacity	Users (and often regulators) cannot observe why an AI system makes the decisions it makes	Standard disclosure and transparency remedies may be inadequate

Athey and Morton argue that standard antitrust frameworks — focused on price effects and market shares — are insufficient for AI markets. The relevant competitive harms may lie in data

accumulation, model quality, and the subtle shaping of user behavior through algorithmic design, none of which traditional tools easily measure or remedy.

2.2 The Dissolution of Firm Boundaries: Shahidi et al.

The theory of the firm, in its Coasian formulation, holds that firms exist where the transaction costs of market coordination exceed the costs of hierarchical organization. AI changes both sides of this calculus simultaneously, with potentially profound implications for firm structure and strategy.

Shahidi and colleagues' 2025 contribution documents how AI is dissolving traditional firm boundaries in three directions. Vertically: AI enables sophisticated supplier relationships that were previously too complex to manage at arm's length, reducing the need for vertical integration. Horizontally: AI allows rapid scaling without proportional increases in human capital, enabling firms to remain much smaller while competing at a large scale. Temporally: AI accelerates the obsolescence of current capabilities, reducing the value of stable, long-term firm-specific human capital.

The result is what Shahidi and colleagues call the platform paradox: the largest AI-enabled firms are simultaneously enormous in economic impact and small in traditional measures (headcount, fixed assets). This creates genuine difficulties for competition policy, labor regulation, and tax administration — all of which were calibrated on the assumption that large economic power implies large, measurable organizational presence.

2.3 AI Agents as Economic Actors: Hadfield and Koh

Perhaps the most conceptually novel contribution in this domain comes from Hadfield and Koh's 2025 paper on AI agents as autonomous economic actors. As AI systems gain the capacity to enter contracts, manage resources, and act on behalf of principals without continuous human oversight, the standard legal and economic framework — which assumes that economic actors are humans or human-controlled entities — begins to break down.

Hadfield and Koh develop a framework for understanding AI agents along two dimensions: the degree of autonomy (from human-supervised to fully independent) and the scope of action (from single-task to broad-domain). Current AI agents — booking systems, trading algorithms, customer service bots — occupy the low-autonomy, narrow-scope corner of this space. But the trajectory is clear: agents are moving toward broader scope and greater autonomy.

The economic implications are significant. When AI agents can execute complex multi-step tasks — negotiate supplier contracts, manage investment portfolios, hire contractors, litigate disputes

— the locus of economic decision-making shifts in ways that have no precedent. Questions of liability, accountability, and the alignment of agent behavior with principal interests become not merely legal technicalities but central economic problems.

2.4 Intelligence as a Commodity: Chatterji, Rock, and Talamás

Chatterji, Rock, and Talamás offer what may be the most unsettling competitive insight in this Part: as AI capabilities become widely accessible through APIs and cloud services, intelligence itself risks becoming a commodity.

In traditional markets, competitive advantage derives from superior resources, capabilities, or information. If AI democratizes access to sophisticated analytical capabilities — allowing any firm with an API key to leverage state-of-the-art prediction, synthesis, and generation — then the competitive advantage previously held by firms with large analytical teams evaporates. What replaces it?

Chatterji and colleagues argue that the answer lies in proprietary data, execution capability, and what they call situated intelligence: the ability to deploy AI in ways that are specific to particular customer relationships, organizational contexts, and institutional knowledge that cannot easily be replicated by a generalist model. The implication for strategists is direct: firms that have treated AI as a capability to acquire (buy the model, hire the engineers) rather than a capability to embed (integrate deeply with proprietary data and workflows) will find their advantage eroded faster than they expect.

PART III

Labor, Distribution, and Human Welfare

No domain of AI economics generates more public concern — or more confused debate — than its implications for workers. The research surveyed here moves beyond simplistic automation-versus-augmentation dichotomies to offer more nuanced frameworks for understanding labor market adjustment, distributional consequences, and ultimately what work means for human flourishing.

3.1 The Adaptive Capacity Index: Manning and Aguirre

Alan Manning and colleagues' 2025 contribution introduces the Adaptive Capacity Index — a framework for evaluating workers' ability to adjust to AI-driven labor market change. Rather than asking which jobs AI will eliminate (a question that yields wildly varying predictions in reliability), they ask which workers are most and least equipped to navigate the transition.

The Adaptive Capacity Index rests on four components: skills transferability (the degree to which current skills apply to adjacent roles); access to retraining (financial, geographic, and institutional); social capital (networks that provide information about and access to new opportunities); and psychological resilience to occupational disruption. The research finds that these factors are highly correlated with existing measures of privilege — age, education, income, geography — meaning that AI-driven labor disruption is likely to compound existing inequalities rather than redistribute opportunity randomly.

Managerial Implication: The Internal Labor Market as Strategic Asset

Manning and Aguirre's framework has a direct organizational implication that is underappreciated in most AI transition planning. Firms that invest in systematically raising the adaptive capacity of their existing workforce — through retraining, transparent career pathway communication, and deliberate skills-bridging — are not merely being generous; they are managing a strategic asset. Workforce adaptive capacity is a source of organizational resilience that cannot easily be purchased on the external labour market.

3.2 Essential Work vs. Accessory Work: Restrepo's Distinction

Pascual Restrepo's 2025 labor economics contribution introduces a distinction that may prove as durable as the canonical skill-biased versus labor-biased technical change framework: the difference between essential work and accessory work.

Essential work refers to tasks that are intrinsic to the production of a good or service — the activities without which the output does not exist or cannot be delivered. Accessory work refers

to tasks that exist only because of the costs and limitations of the technology available — the workarounds, translation layers, and manual overrides that humans perform precisely because automated systems are imperfect.

Restrepo's key empirical finding is that a substantial fraction of white-collar employment — particularly in administrative, compliance, and coordination roles — consists primarily of accessory work: tasks that exist only because enterprise software is not quite good enough to handle exceptions automatically. As AI improves, this accessory work is eliminated not because the underlying activity disappears, but because it no longer requires human execution.

This framework helps explain a phenomenon that puzzles many labor economists: highly educated workers in knowledge-intensive sectors can face significant displacement even when demand for the underlying service (legal advice, financial analysis, medical diagnosis) is growing. The answer is that their roles contained more accessory work than either they or their employers recognized.

Essential Work (more durable)	Accessory Work (AI-vulnerable)
Creative synthesis requiring genuine novelty	Exception handling in automated workflows
Complex stakeholder relationship management	Format translation and data standardization
Judgment calls with significant ethical stakes	Rule-based compliance checking and reporting
Tasks requiring embodied physical presence	Meeting summarization and document routing
New knowledge generation at the frontier	Routine client communication and status updates

3.3 Beyond GDP: Stevenson on Human Meaning

Betsey Stevenson's 2025 contribution represents the most fundamental challenge to standard economic welfare analysis in this volume. The core of her argument: as AI takes over more of the economically measurable work of the economy, we cannot assess its welfare implications through the lens of income and consumption alone.

Work, Stevenson demonstrates through extensive empirical evidence, is not only an input into production and a source of income. It is a primary source of social connection, identity, purpose, and structured engagement with the world. Workers who exit the labor force — whether through unemployment, early retirement, or AI displacement — experience welfare losses that standard income-based measures systematically understate.

> *"We have built an entire apparatus of economic measurement — GDP, productivity, income distribution — on the assumption that the economy's*

> *purpose is consumption. But human beings do not experience their lives primarily as consumers; they experience them as doers, makers, and members of communities. AI forces us to confront this directly."*
> **— Stevenson, 2025**

The practical implication for business leaders is uncomfortable but important. Firms that achieve significant AI-driven productivity gains by reducing headcount may be creating value in the conventional economic sense while simultaneously destroying social value — fracturing communities, eroding the tax base that funds public services, and contributing to the long-run erosion of trust in institutions. These are not merely external costs to be noted in a CSR report; they create political and regulatory risks that are directly material to business strategy.

3.4 Algorithmic Choice Architecture: Ludwig et al.

Ludwig and colleagues' 2025 paper introduces a framework that bridges behavioral economics and AI governance: algorithmic choice architecture. Building on the insights of behavioral economics — that the framing of choices systematically affects decision outcomes — they analyze the unique power of AI systems to shape the information environment in which billions of people make decisions.

Traditional choice architecture (the arrangement of a cafeteria, the default settings of a retirement plan) is relatively static and transparent. AI-driven choice architecture is dynamic, personalized, and opaque. An AI-powered recommendation system does not merely present options; it selects which options to present, in what order, with what framing, and adapts to each individual user's psychological profile.

Ludwig and colleagues find that this creates both enormous potential value — better-aligned choices, improved health and financial behaviors — and significant risks of manipulation and welfare reduction. The regulatory challenge is that the same technical capability that enables beneficial nudges (encouraging saving, exercise, preventive healthcare) can be deployed for extractive purposes (maximizing engagement at the cost of wellbeing, exploiting cognitive biases for commercial gain) with no visible change in system architecture.

PART IV

Information, Knowledge, and Systemic Risks

AI is fundamentally an information technology — it is built from information, it processes information, and it produces information. Understanding its systemic implications for how knowledge is created, validated, and shared is therefore essential to any complete economic analysis.

4.1 The Knowledge Creation Revolution: Mullainathan and Rambachan

Sendhil Mullainathan and Ashesh Rambachan's 2025 paper may be the most epistemically ambitious contribution in the NBER program. Their argument: AI does not merely accelerate existing methods of knowledge creation; it changes the fundamental methods themselves.

Science, since the Enlightenment, has operated through a recognizable methodology: hypothesis generation by human experts, experimental design, data collection, statistical testing, peer review, and gradual consensus formation. This process has been extraordinarily productive but has important structural limitations — it is slow, expensive, and biased toward hypotheses that fit within existing theoretical frameworks (because funding, peer review, and journal acceptance all depend on resonance with existing expert consensus).

AI-enabled knowledge creation breaks from this model in several ways. Large language models can generate novel hypotheses by identifying patterns across literature that are too large for any human to synthesize. Reinforcement learning systems can explore solution spaces that humans would never enter because they are too counterintuitive to propose as hypotheses. AI simulation can test causal theories at scales and speeds impossible in physical experiments.

But Mullainathan and Rambachan are careful to note the epistemological risks alongside the opportunities. AI systems can identify spurious correlations, generate plausible but wrong hypotheses, and — most dangerously — produce confident-seeming outputs that are systematically biased by training data that reflects the prejudices and blind spots of the human knowledge it learned from. The risk is not that AI will know too little, but that it will be confidently wrong in ways that are difficult to detect.

4.2 Hayek's Challenge Revisited: Brynjolfsson and Hitzig

Erik Brynjolfsson and Zoe Hitzig's 2025 paper takes as its starting point one of the most important arguments in twentieth-century economics: Hayek's 1945 insight that the market price system is

the most efficient known mechanism for aggregating dispersed local knowledge — knowledge that no central planner could ever collect, process, and act upon.

Brynjolfsson and Hitzig ask: Does AI change this? Can sufficiently powerful AI systems process enough information to replicate or exceed the knowledge-aggregation function of decentralized markets?

Their answer is nuanced and important. On the aggregation of explicit, numerical information — market prices, demand quantities, production capacities — AI may indeed approach or exceed market efficiency under certain conditions. But Hayek's deeper point was about tacit knowledge: the kind of knowledge that cannot easily be articulated, that exists in practices, routines, and relationships, and that is lost when it is translated into data. This kind of knowledge, Brynjolfsson and Hitzig argue, remains stubbornly resistant to AI capture — and it is precisely this knowledge that underlies much of the most important human economic activity.

The policy implication is significant. Proposals for AI-enabled central planning or algorithmic market design should be evaluated with Hayek's challenge in mind: the question is not whether AI can process more information than human markets, but whether it can access the right kinds of information. In most domains of complex economic activity, it cannot — at least not yet.

4.3 The Information Ecosystem at Risk: Stiglitz and Ventura-Bolet

Joseph Stiglitz and colleagues' 2025 contribution sounds a note of warning that sits in productive tension with the more optimistic readings of AI's informational potential. Their argument: the same AI capabilities that improve information processing in many domains can simultaneously degrade the broader information ecosystem on which economic and political decision-making depends.

The mechanism is straightforward but underappreciated. The economic value of information is partly absolute (knowing something true is valuable) and partly positional (knowing something others do not know is valuable). AI dramatically reduces the cost of generating plausible-seeming content — including plausible-seeming but false content. In an environment of abundant AI-generated text, video, and analysis, the cost of creating disinformation falls while the cost of verifying authenticity rises.

Stiglitz and Ventura-Bolet model this as a market failure: the private returns to generating AI content are high (attention, engagement, commercial influence) while the social costs — the degradation of shared epistemic standards, the erosion of trust in institutions, the increased

difficulty of coordinating on true beliefs — are dispersed and do not flow back to the content generators. The result is a classic negative externality that standard market mechanisms will not correct without intervention.

Strategic Implication: Epistemic Risk as Business Risk

The Stiglitz and Ventura-Bolet framework has direct implications for business leaders that are not always appreciated. As the information ecosystem degrades, the cost of all forms of market intelligence, due diligence, and stakeholder communication rises. Firms whose business models depend on consumer trust (financial services, healthcare, media) face particular exposure. The degradation of shared epistemic standards is not an abstract political problem; it is a business environment risk that deserves explicit treatment in enterprise risk management frameworks.

PART V

Policy Responses and Long-Term Considerations

The final Part addresses the most difficult questions: how should fiscal policy adapt to an economy that AI is making increasingly unmeasurable, and how should society weigh the potential benefits of transformative AI against the possibility of catastrophic harm?

5.1 The Fiscal Challenge: Korinek and Lockwood

The fiscal implications of AI transformation are both technically complex and politically combustible. Korinek and Lockwood's 2025 analysis provides the most systematic treatment available of how AI erodes the traditional foundations of public finance.

Modern fiscal systems rest on three interconnected tax bases: income earned by workers (personal income and payroll taxes), profits earned by firms (corporate income tax), and consumption (sales and value-added taxes). Each of these bases is threatened by AI in distinct ways.

- Labor income taxes: as AI displaces workers from paid employment, the payroll tax base shrinks. This is not merely a cyclical adjustment; if AI-driven displacement is structural and permanent in certain occupations, the long-run revenue base for social insurance programs (pensions, unemployment insurance, healthcare) may be fundamentally inadequate.
- Corporate income taxes: as AI shifts value creation toward intangibles, data assets, and cross-border digital services, the geographic and legal attribution of corporate income becomes increasingly arbitrary. The existing corporate tax framework was built around physical assets and measurable production; it is poorly equipped for an economy where value is created by AI models trained on data collected globally, operated from tax-favorable jurisdictions, and consumed everywhere.
- Consumption taxes: paradoxically, consumption taxes may be the most resilient, since humans must eventually consume in the physical world. But as AI enables sophisticated tax arbitrage and blurs the line between consumption and investment, even this base is not immune.

Korinek and Lockwood do not offer a fully worked-out alternative fiscal architecture — they are appropriately honest about the difficulty of designing robust tax systems for an economy they cannot fully measure. But they identify several design principles: tax bases should shift toward sources of value that cannot easily relocate or become immeasurable (land and natural resources, energy consumption, physical transactions); taxation of AI capabilities themselves — robot taxes, data taxes, automation levies — should be evaluated for their incentive effects; and

international coordination on AI taxation is essential if jurisdiction-shopping is not to hollow out national tax bases.

5.2 The Existential Question: Chad Jones on Catastrophic Risk

The final and most profound set of questions addressed in the NBER program concerns not AI's distributional or fiscal implications, but its potential to generate catastrophic or existential risks: outcomes so bad that no amount of economic benefit could compensate for them.

Chad Jones' 2025 contribution applies the tools of growth theory and cost-benefit analysis to perhaps the most uncomfortable question in contemporary economics: how much should society spend to reduce the probability of an AI-related catastrophe?

Jones begins with an observation that grounds the analysis: in standard economic frameworks, we accept very large expenditures to reduce small probabilities of catastrophic outcomes. We build nuclear power plants to survive earthquakes of very low probability. We invest in pandemic preparedness for once-in-a-century events. The question for AI safety is whether the expected costs of AI catastrophe — properly accounting for the magnitude of potential harm, its irreversibility, and its probability — justify equivalent or greater investment.

> *"The tools of economics are not ill-suited to existential risk. They are, in fact, precisely the tools we need: we are making decisions under uncertainty about outcomes of vastly different magnitudes. The question is not whether to apply cost-benefit thinking, but how to do so rigorously."*
> **— Chad Jones, 2025**

Jones develops a framework for thinking about AI safety investment that incorporates four factors: the probability of catastrophic outcomes under different development scenarios; the magnitude and irreversibility of potential harm; the degree to which safety investment reduces risk (the 'safety production function'); and the opportunity cost of diverting resources from beneficial AI development.

His central finding — deliberately stated as a range rather than a point estimate, given deep uncertainty — is that even relatively modest estimates of catastrophic risk probability justify safety expenditure at a scale that current industry and government investment falls far short of. The intuition is Pascalian: when the potential harm is civilizational in magnitude, even low probabilities imply enormous expected costs.

For business leaders, this framework matters in two ways. First, it provides a conceptual grounding for why AI safety is not merely a regulatory compliance issue or a reputational risk to be managed, but a genuine economic priority that should be reflected in investment decisions. Second, it suggests that firms operating at the frontier of AI capability have a particular responsibility — and a particular exposure — to the costs of inadequate safety investment, both in terms of direct liability and in terms of the political and regulatory responses that a major AI-related disaster would provoke.

Synthesis: An Integrated Framework for AI Strategy

The literature collectively suggests that AI transformation is not merely a technological disruption. It is a transition between economic paradigms.

The emerging AI economy may feature:

Industrial Economy	AI Economy
Scarce intelligence	Abundant intelligence
Scarce information	Scarce trust
Stable firms	Fluid ecosystems
Labor-centric taxation	Intangible capital taxation
Human cognitive advantage	Human relational advantage
Productivity focus	Meaning and welfare focus
Observable markets	Hidden digital activity
Human-only agency	Human-AI hybrid agency

Having surveyed the five domains of AI economics, we can now assemble an integrated framework for how business leaders and policymakers should think about AI transformation. The framework rests on four propositions.

Proposition 1: AI is a Measurement Problem Before It Is a Management Problem

Every domain surveyed in this paper — from innovation to labor markets, from fiscal policy to systemic risk — encounters the same foundational obstacle: our measurement systems were not designed for the economy AI is creating. Firms that build internal capability to measure what actually matters — not just financial performance, but also capability accumulation, workforce adaptive capacity, information ecosystem health, and AI safety posture — will have a systematic advantage over those that rely on standard accounting and reporting frameworks.

Proposition 2: The Prediction-Judgment Frontier Is the Central Strategic Variable

The Agrawal-Gans-Goldfarb framework of prediction becoming cheap while judgment becomes scarce provides the most practically useful strategic lens in this literature. Every organization should be able to answer: where on this frontier does our competitive advantage currently sit? How is AI moving that frontier? What investments in human judgment capability are we making to stay ahead of it? These questions are not abstract; they should be embedded in every major strategic planning cycle.

Proposition 3: Distributional Consequences Are Strategic Risks, Not Just Ethical Concerns

The labor economics literature — Manning, Restrepo, Stevenson — consistently demonstrates that AI-driven disruption has distributional consequences that compound existing inequalities and generate social and political dynamics that directly affect the business environment. The collapse of the middle-skill labor market, the degradation of communities whose economies depended on displaced occupations, and the erosion of the tax base that funds public services — these are not side effects of AI transformation that firms can externalize indefinitely. They generate regulatory responses, political instability, and shifts in consumer sentiment that are material to long-term business strategy.

Proposition 4: The Transition Dynamics Are as Important as the Long-Run Equilibrium

Much AI optimism is explicitly long-run: in the long run, AI will generate enormous wealth, new forms of work will emerge, and social institutions will adapt. This may well be true. But economies and societies do not live in the long run; they live in the transition. History suggests that transitions driven by General Purpose Technologies are marked by decades of disruption, institutional strain, and distributional conflict before the new equilibrium stabilizes. The economic frameworks surveyed here are most practically valuable not as guides to the eventual destination, but as tools for navigating the path.

Domain	Key Authors	Core Finding	Practitioner Implication
Growth & Innovation	Ben Jones; Agrawal et al.	Prediction costs fall; judgment becomes scarce; super linear innovation possible if AI relieves bottlenecks	Map prediction/judgment frontier; invest in situated intelligence
Measurement & Accounting	Coyle/Poquiz; Korinek/Lockwood	Standard metrics undercount AI value; tax bases erode with intangibles	Build internal measurement for capability and adaptive capacity
Competition & Organization	Athey/Morton; Chatterji et al.; Hadfield/Koh	Non-rival AI creates natural monopolies; agents blur firm/market line; intelligence commoditizes	Prioritise proprietary data and embedded AI over bought capability
Labor & Distribution	Manning; Restrepo; Stevenson	Accessory work displaced; adaptive capacity compounds inequality; work has welfare value beyond income	Raise workforce adaptive capacity; treat labor transition as strategic risk
Information & Knowledge	Brynjolfsson/Hitzig; Mullainathan; Stiglitz et al.	AI changes knowledge creation but tacit knowledge persists; disinformation degrades ecosystem	Treat epistemic risk in ERM; invest in verification and trust infrastructure
Policy & Governance	Korinek/Lockwood; Chad Jones	AI erodes tax bases; safety expenditure justified by expected catastrophic harm	Engage in international tax coordination; treat AI safety as investment, not compliance

Conclusion: Managing the Most Consequential Transition in Economic History

The economic scholarship surveyed in this paper does not offer easy comfort. The transition ahead is genuinely hard: it involves measurement failures that will obscure what is happening, distributional tensions that will generate political conflict, institutional structures that are poorly calibrated for the economy AI is creating, and — at the limit — risks of harm severe enough to warrant serious, sustained investment in prevention.

But the same body of scholarship offers something more valuable than comfort: it offers intellectual frameworks adequate to the scale of the challenge. The prediction-judgment frontier gives strategists a language for mapping competitive dynamics. The adaptive capacity framework provides HR leaders with a tool to assess workforce resilience. The measurement gap analysis gives CFOs a reason to invest in internal metrics that go beyond GAAP. The catastrophic risk framework gives boards a justification for safety investment that can withstand scrutiny.

What the research does not offer — and is honest about not offering — is certainty about the pace or ultimate destination of AI transformation. The scholars of the NBER program have done the intellectual work of developing frameworks rigorous enough to be useful without pretending to a precision that the uncertainty of the moment does not allow.

That epistemological humility is itself the most important lesson for practitioners. The leaders who will navigate AI transformation best are not those who claim to know exactly what is coming, but those who have built the analytical frameworks and organizational capabilities to adapt rapidly as the picture becomes clearer. In the economics of transformative AI, as in all strategic planning under genuine uncertainty, the ability to update is more valuable than the accuracy of any initial forecast.

The work of economic scholarship is never finished, but the work of economic statecraft cannot wait for it to be.

— Harvard Business School, Strategy & Competitiveness, 2025

Source Reference Guide

The following is a structured guide to the primary sources synthesised in this paper. Readers are encouraged to consult the original works for technical depth and empirical detail.

Author(s)	Year	Work	Core Contribution
Agrawal, Gans, Goldfarb	2019, 2025	NBER volume + 2025 update on 'genius on demand'	Prediction-judgment frontier; AI as reduction in prediction cost
Brynjolfsson, Korinek, Agrawal	2025	A Research Agenda for the Economics of Transformative AI	Comprehensive research framework; overarching synthesis
Ben Jones	2025	Innovation process and bottleneck framework	Rising knowledge burden; AI as bottleneck relief; superlinear growth
Coyle and Poquiz	2025	AI measurement gap	Statistical systems' inability to measure AI value creation
Korinek and Lockwood	2025	AI and tax base erosion	Fiscal implications of unmeasurable AI value; tax reform principles
Athey and Morton	2025	AI's unique competitive economics	Non-rivalry, feedback loops, data network effects in AI markets
Hadfield and Koh	2025	AI agents as autonomous economic actors	Autonomy-scope framework; liability and accountability challenges
Shahidi et al.	2025	Dissolution of firm boundaries	Platform paradox; vertical/horizontal/temporal firm restructuring
Chatterji, Rock, Talamás	2025	Intelligence as commodity	Commoditization of AI; situated intelligence as competitive moat
Manning and Aguirre	2025	Workers' adaptive capacity index	Four-component resilience framework; inequality amplification
Restrepo	2025	Essential vs. accessory work	Structural basis of white-collar displacement; task decomposition
Stevenson	2025	Work and human meaning	Beyond GDP welfare analysis; work as identity and social connection

Ludwig et al.	2025	Algorithmic choice architecture	AI-driven nudges; welfare benefits and manipulation risks
Brynjolfsson and Hitzig	2025	Hayek challenge and decentralized knowledge	Limits of AI central planning; tacit knowledge persistence
Mullainathan and Rambachan	2025	AI and knowledge creation methods	Epistemological transformation of science; confident error risks
Stiglitz and Ventura-Bolet	2025	AI and information ecosystem	Disinformation externality; epistemic market failure
Chad Jones	2025	Existential risk cost-benefit analysis	Economic framework for AI safety investment; catastrophic risk pricing
Daniel Kahneman	2017	NBER inaugural AI conference address	Prediction, decision-making, and the cognitive economics of AI

Discussion Questions for Practitioners

- How can firms measure the intangible value created by AI?
- What organizational structures will thrive in an AI-augmented economy?
- How should governments tax AI-driven rents without stifling innovation?
- What reskilling programs are most effective for workers displaced by AI?
- How can we preserve trust in an era of AI-generated information?