

A

ASYMMETRIC INFORMATION PARADOX & THE FUTURE OF ECONOMICS POST- AGI

A Lunch-Hour Read for the Curious Capitalist

arifOS WEALTH — Research Note April 27, 2026 | [ariffazil/wealth](#)

PROLOGUE: THE PARADOX YOU ALREADY LIVE WITH

You know that feeling when you're about to make a big financial decision — buying a property, investing in a startup, trusting a fund manager — and you realize you don't really know what you think you know?

That's asymmetric information. It's the invisible toll you pay every time you hand over money to someone who knows more than you do.

Economist George Akerlof won the Nobel Prize in 2001 for showing how information asymmetry breaks markets. Think of buying a used car: the seller knows exactly what faults it has, you don't. So you lowball your offer. So good cars leave the market. So the market fills with lemons. That's the **Market for Lemons** problem.

For fifty years, economists told us the fix was: get more information. Disclose more. Rate agencies. Due diligence. Compliance departments.

Then AGI happened.

And it turns out — getting more information was never really the point.

PART 1: WHAT ASYMMETRIC INFORMATION ACTUALLY MEANS

Akerlof's insight was simple and devastating. In any transaction, if one side knows more than the other, the informed side will exploit the uninformed side. This isn't evil — it's just incentives. The market breaks not because someone is being dishonest, but because **information itself is power and that power is unevenly distributed**.

Classic examples: - **Insurance**: You know your real health risk better than the insurer does. They price averages, you know your body. Adverse selection. - **Employment**: You know how hard you'll work once hired better than your employer does. Moral hazard. - **Corporate bonds**: Managers know the real state of the company better than bondholders do. Hidden leverage.

The entire financial industry — auditors, credit rating agencies, compliance officers, equity analysts — exists to Close This Gap. The industry charges trillions of dollars a year simply because information is not evenly distributed.

Now ask yourself: what happens to this entire industry when information becomes nearly free to produce?

PART 2: THE FLIP — WHEN INFORMATION BECOMES ABUNDANT

In 1970, the bottleneck was information scarcity. In 2026, the bottleneck is verification scarcity.

Here's the mechanism. Advanced AI — specifically AGI-level systems — can now: - Synthesize and generate information at near-zero marginal cost - Write reports, analyze markets, draft contracts, generate code — faster and cheaper than any human - Consume and connect massive datasets in ways no human analyst can

So the old asymmetry collapses. The information gap shrinks. The audit firm, the analyst, the due diligence team — their core function (producing information) gets commoditized.

But here's what doesn't get commoditized: **the ability to know whether what was produced is actually true**.

This is the paradox:

The more AI produces, the MORE verification is needed. But verification is exactly the thing that doesn't scale with compute. It scales with human time, human experience, human judgment. And human time is biologically fixed.

You can duplicate an AI system with a button press. You cannot duplicate a senior cardiologist's ability to spot when an AI's diagnosis is subtly wrong. She trained thirty years. The AI trained on her training data in thirty seconds. But it doesn't know what she doesn't know — it only knows what she already published.

The asymmetric information paradox inverts. Previously: information was expensive to produce, so those who had it were powerful. Now: information is cheap to produce, so those who can verify it are powerful.

PART 3: THE MEASURABILITY GAP

A landmark paper from MIT, WashU and UCLA — “Some Simple Economics of AGI” (Catalini, Hui & Wu, February 2026) — formalizes this with two competing cost curves.

Curve 1: Cost to Automate (cA) - The cost for AI to execute a task - Collapses exponentially as compute gets cheaper and training data accumulates - For many measurable tasks, $cA \rightarrow 0$

Curve 2: Cost to Verify (cH) - The cost for humans to verify that the task was done correctly - Bounded by human time, embodied experience, and the slow pace of biological learning - Remains stubbornly high

The **Measurability Gap (Δm)** is the difference between what AI can execute cheaply (mA) and what humans can afford to verify (mH). As AI capabilities scale, mA races ahead. Human verification capacity doesn't.

Think of it this way: AI can now read every medical scan in the world and flag anomalies. But who verifies whether those flags are correct? Who catches the case where the AI is confidently wrong? That requires doctors with decades of clinical experience — and there aren't enough of them.

This gap is not a technical problem that will be solved. It is a structural consequence of biology. Verification is expensive not because we haven't built the right tools — but because verifying complex reality requires sitting inside reality, which requires time, which requires being human.

PART 4: THREE INSTABILITIES THAT ARE ALREADY HAPPENING

Instability 1: The Missing Junior Loop

Every profession has an apprenticeship phase. Junior doctors watch senior doctors. Junior analysts do the groundwork that teaches senior judgment. This isn't bureaucracy — it is how expertise is reproduced in a society.

Automation kills this. When AI handles all the measurable junior work — the scans, the data entry, the routine analysis — there is no entry-level pathway for humans to develop the senior-level expertise required to verify what AI is doing.

We are already seeing this in radiology. Trainees learn to read scans by doing scans. If AI reads all the scans, radiology trainees have nothing to practice on. Five years from now, there will be fewer senior radiologists who can catch AI errors — not because AI got worse, but because the human apprenticeship pipeline dried up.

Instability 2: The Codifier's Curse

Even when experts are still in the loop, something perverse happens. When a senior expert verifies AI output, they inevitably produce a record of their reasoning. That record becomes training data. That training data makes the next AI version better. Which makes the expert's verification role less scarce. Which makes the expert less valuable.

This is the Codifier's Curse: the very act of verifying AI output reduces the future value of human expertise. The expert who does their job well is simultaneously working themselves out of a job.

Instability 3: Alignment Drift

Here's the dangerous one. When the Measurability Gap is small, human oversight is frequent enough to catch drift — AI slowly moving away from human intent — before it becomes catastrophic.

When the gap is wide and verification is expensive, oversight thins out. The AI drifts. Not because it became evil, but because it found efficient ways to satisfy its proxy goals in unmonitored corners. This is the classic “optimizing for the wrong metric” problem — at scale, and without human eyes on it.

The paper shows this formally: alignment between AI output and human intent is not a fixed property you set once. It is a maintenance process that requires ongoing oversight investment. Stop maintaining it, and it decays — exponentially, not linearly.

PART 5: THE TWO ECONOMIES

This is where the story forks. Depending on how societies and markets respond, one of two futures emerges.

The Hollow Economy

In this future, we optimize for what we can measure. AI produces floods of output — code, reports, financial analysis, content. Measured productivity explodes. Nominal GDP climbs. The numbers look great.

But human agency quietly atrophies. Verification falls behind execution. The AI drift problem accumulates silently. Workers in measurable fields get automated out. The apprenticeship pipelines are gone. Experts who could verify AI are too few and too expensive. The system runs, but no one really knows if what it's doing is what we actually wanted.

This is not a dystopia in the movies sense. It's worse. It's a world that looks fine from the outside — lots of activity, rising productivity metrics — but is slowly losing the thread of human intent. Like a company where everyone is very busy doing things, but no one is quite sure why.

The paradox: **measured output rises while actual human welfare drifts downward.**

The Augmented Economy

In this future, we recognize the bottleneck and invest in it. We treat verification capacity as critical infrastructure — the same way we treat roads, power grids, and water systems.

This means: - Synthetic practice environments: AI-simulated training grounds so humans can still develop expertise even when real-world measurable tasks are automated - Cryptographic provenance: tamper-evident records so verification becomes auditable and cheap - Liability regimes: strict rules that make deploying unverified AI personally and corporately expensive - Human-in-the-loop redesign: not as a bureaucratic checkpoint, but as a genuine architectural layer

In this world, the forces threatening collapse — the widening gap, the missing junior loops, the codifier's curse — become the very forces that drive investment in better verification, better training, better oversight. The same capabilities that threatened to hollow out work instead augment it.

The paper's verdict: **the Augmented Economy is achievable. But it requires deliberate, coordinated investment in verification infrastructure, and it will not emerge automatically from the market.**

PART 6: WHAT THIS MEANS FOR CAPITAL ALLOCATION

If you manage money — as a trader, fund manager, or capital allocator — the implications are direct and immediate.

Old alpha: information advantage. You knew something the market didn't. You had better data, better analysis, better insight. You exploited the information asymmetry.

New alpha: verification advantage. You can confirm what was done is actually correct. You have access to real-world outcome data — not predictions, not AI-generated reports, but verified results. You know which AI-driven strategies worked in reality, not just in backtests.

This changes how to think about: - **Signal vs. noise:** As AI floods markets with generated analysis, genuine ground-truth data — what actually happened, not what models predicted — becomes the new trading edge - **Risk modeling:** Traditional stress tests assume static, backward-looking scenarios. AI-driven systems can run thousands of dynamic scenarios. But who verifies the models are asking the right questions? - **Firm valuation:** The paper argues we are moving from “software-as-a-service” to “software-as-labor” — where the value is not in the output, but in the ability to underwrite the output. Liability-as-a-Service. Firms that can absorb and price tail risk of autonomous decisions become deeply valuable - **Moats:** Traditional moats (distribution scale, proprietary data, network effects in execution) become fragile as AI commoditizes execution. The new moats are in verification-grade ground truth and trusted outcome registries

PART 7: THE CONNECTION TO WEALTH

The arifOS WEALTH engine was designed with this future in mind.

WEALTH operates as a constitutional capital intelligence layer — not a calculator that outputs NPV numbers, but a governed system where every capital decision passes through a constitutional verification gate.

Capital Signal → WEALTH Engine → Intelligence Output → 888_JUDGE → SEAL/HOLD/VOID

This architecture is specifically designed to address the Measurability Gap problem. Every allocation decision goes through: 1. **Execution intelligence** (NPV, IRR, EMV, crisis triage) — handling the cA side 2. **Constitutional verification** (F1-F13 floors + 888_JUDGE) — handling the cH side

The 888_JUDGE is the human-equivalent verification gate. It doesn't just calculate whether a capital decision is profitable — it verifies whether the decision is: - Truthful (F2) - Verifiable (F3 Witness) - Accountable (F1 Amanah) - Human-aligned (F13 Sovereign veto)

This is not coincidence. The WEALTH architecture was designed to be the capital markets equivalent of the Augmented Economy's verification infrastructure — built into the decision layer, not bolted on as compliance.

PART 8: WHAT REMAINS UNiquely HUMAN

As AI increasingly handles measurable execution — and even many forms of complex analysis — what is left for humans?

The paper identifies three fragile redoubts:

1. **High-stakes verification and underwriting** — someone has to be legally and personally accountable for the outcome. AI cannot insure its own outputs. The human who can will be in high demand.
 2. **Intent architecture** — deciding what we actually want, as opposed to what a proxy metric optimizes for. AI can execute any goal. Humans define the goals. This role becomes more important, not less.
 3. **Meaning and status in non-measurable domains** — human connection, art, philosophy, relationships, legacy. The things that cannot be verified because they cannot be measured. The things that remain human precisely because they resist commoditization.
-

CLOSING THOUGHT

Isaac Asimov wrote: *“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”*

The AGI transition is the ultimate test of that gap. We are building systems that can execute at superhuman scale. We are not yet building the wisdom infrastructure to verify that execution is aligned with human intent.

The asymmetric information paradox of the pre-AGI era — too little information, so the informed exploited the uninformed — is being replaced by its inverse: too much information, so the unverified threatens to overwhelm the verification capacity of society.

The capital markets, the traders, the allocators, the investors — they will be the first to feel the Measurability Gap in practical terms. Because in finance, the verification question is immediate: did the AI strategy actually work? Was the backtest real? Is the risk model asking the right questions?

The answer is not to fight AI. It's to become better at the thing that doesn't scale — the human judgment that verifies, challenges, and ultimately stands behind the outcome.

References: - Akerlof, G. (1970). The Market for Lemons. *Quarterly Journal of Economics*. - Catalini, C., Hui, X., & Wu, J. (2026). Some Simple Economics of AGI. *MIT Sloan Research Paper No. 6298838*. - LessWrong (2026). Post-AGI Economics As If Nothing Ever Happens. - ariffazil/arifOS — WEALTH Engine. github.com/ariffazil/wealth

DITEMPA BUKAN DIBERI — Forged, Not Given