

Artificial Intelligence and Economic Growth: A few finger-exercises

[A response to a discussion between
Robin Hanson and Eliezer Yudkowsky
at [Overcoming Bias](#)]

Garett Jones
George Mason University

January 2009

One way to see how Artificial Intelligence (AI) might impact the economy would be to follow Robin's model of "test near, apply far." Chad Jones ([Fig. 1, p. 763](#), and in his short, readable text *Introduction to Economic Growth*) has reminded economists that the number of scientists and researchers has more than doubled in the G-5 countries since 1950, while the growth rate of living standards hasn't budged: Twice the researchers, zero effect on growth. This fact has major implications for how we think about the effect of AI on economic growth.

I'll assert that progress with computers, lab science, etc. has made the average scientist today twice as productive as the average scientist in 1950. So we can conservatively say that there are four times the number of "quality-weighted" scientists as in 1950. That means that adjusting for increasing quality makes the puzzle of "more science but no more growth" bigger, not smaller.

Clearly, a move to AI would increase the number of quality-weighted scientists by many, many orders of magnitude. But if four times zero is zero, then four hundred times zero is still zero, and four million times zero is still zero.

A lot depends on that zero, of course: If the post-WWII quadrupling of research effort actually raised growth by 0.1% per year (an effect that would be barely detectible using conventional macroeconomic measures), then an AI that multiplied today's annual research efforts by four million would create a singularity.

Models

So, where can we go for metaphors that would help us distinguish between "zero" and "really close to zero?" I think that the modern literature on economic growth and the [economics of innovation](#) has provided a rich set of metaphors that help me think about the economics consequences of AI.

Let me start with a simple metaphor from Weil's undergraduate textbook *Economic Growth*. Weil says that maybe there are two kinds of goods people want, bread and cheese, and that people always consume them in fixed proportions: 2 slices of bread and one slice of cheese. Now suppose that cheese productivity skyrockets—the Enterprise's replicator is really good at making cheese—but bread productivity is stagnant. Well, you'll stop using humans to make cheese, which means those humans can shift over to making bread. Perhaps that will double living standards. But the fact that every slice of cheese needs a pair of inefficiently-produced bread slices means that all that extra potential cheese is worthless. Big innovation boost in one sector has minimal impact on economic growth...sound familiar? It should, since it's a standard pessimist's trope about the irrelevance of AI. Let's probe this metaphor a bit more....

Bread

So, is there some feature of the real-world economy that sounds like bread? Would AI-produced output have to be combined with some scarce, inefficiently-produced output in order to become useful to humans? Eliezer touched on this when he said noted that there might be a smooth transition to a singularity, due to “the difficulties of *design* and *engineering*, even after all manufacturing is solved.” To “design and engineering,” one might add a number of other possibilities.

Human contact comes to mind: Marketing, advertising, glad-handing, trust-building: All the processes that come between the inventor and the customer. Now, if AI can eliminate many of those gaps (e.g., by better simulating what will please customers), then that could eliminate some important [barriers to riches](#).

So, to the extent that there's something like a need for perfect complements (a.k.a., fixed proportions) on either the supply side or the demand side, (i.e., a Leontief production function or Leontief preferences in consumption), we're stuck with rapid innovation causing only modest economic change. As so often in economics, shapes matter: If the shape of the production function is $GDP = \max[\text{ideas}, \text{bodies}]$, then singularities are unlikely.

FOOM for 10% of GDP

A moment more on the demand side: I should emphasize the possibility that people just won't want to spend all their money on machine-made *stuff*. Goodness knows we don't do it now. A truism of macroeconomics is that most of GDP is in the service sector, with the FIRE sector (finance, insurance, real estate) taking about 1/4 of the total. And as Deirdre McCloskey once noted in an interesting talk of the same name, “One quarter of GDP is persuasion.” There's something about humans that makes us need (or want) to spend money on stuff that isn't machine-made.

And that has big consequences. For some strange reason, we measure our spending by what fraction of our *dollars we spend* on this or that category rather than the fraction of *welfare we receive* from products we buy that category. Certainly, if we measured our spending in quality-adjusted terms, there'd be massive changes over time. In 1900, people spent about 1% of GDP on health care, and now it's about 15% in dollar terms. But in quality-adjusted terms, aren't we closer to getting 50% or 90% of our each year's value from health care, a field with major technological progress? And while we're probably spending a greater share of our income on entertainment, aren't we getting vastly higher shares of our value from entertainment? So areas of spending that cost us a small fraction of our total income probably give us a large fraction of our actual welfare.

How does this relate to the singularity? If technological progress in some sectors becomes superexponential, perhaps people will still only spend 15% of income on health care, a few percent more on entertainment, maybe a little on singularity-driven clothing and transportation, but after that, they'll still spend half of it on in-person dining, in-person haircuts, in-person visits to friends. Those human interactions, which remain a stubbornly high percentage of GDP, could easily continue to chew up most of our spending.

Another truism in economics is that manufacturing productivity growth is high, yet manufacturing is still a small and probably shrinking percentage of world GDP. Most singularity discussions turn around questions of manufacturing (though they could easily turn to question of literature and moviemaking). A singularity in manufacturing just might not matter that much to most people. Maybe the manufacturing singularity will be succeeded by a few more singularities that might get us to singularities that cover 90% or more of GDP—then it would surely feel like a different kind of humanity. If Robin's estimates are about right, then singularities should be coming along much faster after this next one....

Two-thirds

One lesson to draw from this discussion is that changing the shape of the production function (or perhaps the utility function) is crucial to making singularities possible. So once you start thinking in terms of the production function metaphor, you can think about this question: "How do we make humans and ideas much more substitutable than humans and machines have been over the last hundred years?"

One of the great puzzles of economics is the surprising fact that the share of income going to workers has remained roughly constant at $2/3$ across decades and across countries over the last century. Economists take that as evidence that capital and labor are only mediocre substitutes for each other, and that the "production function" follows something like a Cobb-Douglas form. There's a vast literature on this production function, and my reading of the literature is that the $2/3$ value tells us something deep about how humans produce stuff with machines. There were endless prophecies that

machines would destroy human employment and that wages would collapse and machines would take the jobs. *Ex ante*, those seem like entirely reasonable prophecies. But that just didn't happen: Might it not happen again?

The endogenous growth literature has tried to come up with metaphors of idea-building that are as thought-provoking and data-reducing as the Cobb-Douglas production function.¹ The best attempt, in my view, is the semi-endogenous growth model of Chad Jones. He presents a fairly complete version of this model in his advanced undergrad text, *Introduction to Economic Growth*. He assumes that the stock of useful ideas grows like this:

$$dA/dt = wL^a A^b$$

Capital letters are variables that change over time, and lower-case letters are parameters whose sizes are up for discussion. Where A is the stock of useful ideas (and dA/dt is the time derivative), L is the number of researchers, w is a parameter that stands for how quickly new ideas come along given all the other values. a and b deserve further discussion. a is usually thought to be weakly between zero and one: If it's one, twice the number of workers yield twice the number of ideas. But in the real world, there's a "stepping on toes" effect, as Chad Jones notes: Two researchers might find the same idea simultaneously, a wasted effort. Also, at a given point in time, diminishing returns could be important.

The b value is usually thought to be weakly less than one. If the number is between zero and one, it's a "standing on the shoulders of giants" effect, in Jones's words. If it's less than zero, it's a "fishing out" effect: We discovered the best ideas early on (Newtonian physics, Lavoisier's discoveries, and *The Origin of Species*, in both chronological order and possibly in order of earth-shatteringness). As time goes on, it'll take ever-increasing effort to find ideas of equal value, if b is negative. But even if b is positive, any value less than one means there's a kind of diminishing returns at work.

Note that you might agree with particular values of a or b , or you might disagree with them: But notice that you can actually have an opinion about a number, and as we'll see in a moment, the values of a and b have sharp predictions for long-run growth and innovation.

¹ Eliezer mentioned some of the combinatorics literature, one strand of the endogenous growth literature, and along the way he noted that one assumption of endogenous growth models is that "there's an unlimited number of really good ideas out there." He seemed to disagree with that idea, noting that "I have friends in venture capital who would laugh like hell at the notion..." But isn't the promise of AI that there really *are* an essentially unlimited number of good ideas, and that the real barrier to riches is that it currently takes a lot of time to search through the alternatives? "Panning for gold" is one of the standard endogenous growth metaphors: It takes time and effort to sort between junk and value. "Recipes" is another metaphor: Same small set of ingredients, many different possible products, most of which are worthless. [Paul Romer's encyclopedia article](#) on economic growth is assigned reading in all my growth courses. It shows the power of combinatorics and indirectly illustrates the power of the "faster panning" that AI could make possible in the future—and that IQ-type intelligence makes possible right now.

Also: The fact that combinatoric potential has been with us, yet frontier-country growth rates have stayed stable over the last century, is evidence that Something Else is holding us back—there's a narrow pipe out there somewhere. What will widen the pipe?

In this essay, as in most textbook treatments, I'll focus on steady states. Let's focus on steady states where the growth rate of technology, $(dA/dt)/A$, is a constant number, since that's what's been true for decades at a time (i.e., it happens in the real world: It was close to zero for millennia, and then it rose to maybe 0.5% per year for a century, then it rose to maybe 2% for a century). Taking logs and time derivatives and simplifying, this yields:

$$\begin{aligned} (dA/dt)/A &= \text{steady-state growth of technology} \\ &= (a*n)/(1-b) \end{aligned}$$

Where n is the growth rate in the number of researchers. So to grow fast in the very long run, you need less "stepping on toes," a higher growth rate (not level!) of the number of researchers, and you need researchers who can see farther with a given stock of knowledge.

Once nice thing about this model is that the steady-state equation captures both the Hanson and Yudkowsky channels of AI-driven economic growth. First, the Hanson channel: If robots become effective replacements for human researchers, then it's easy to imagine an annual growth rate in the number of researchers (n) that is enormous and ever-increasing.

Second, the Yudkowsky channel: AI would be less likely to step on toes and more likely to see farther on the shoulders of giants. So both a and b rise. One implication of this semi-endogenous growth model is that if $b > 1$, then growth becomes unboundedly high and no steady state exists. So that's one way to get a singularity, but not the only one.

This brings me to **Request #1**: I would love to see Eliezer and Robin shoehorn their ideas about the productivity-enhancing prospects of friendly AI into the admittedly ill-fitting semi-endogenous growth literature. I think it would help the conversation if they took positions on the likely values of a , n , and b in a world that had the potential to enter or had just entered a singularity. One might say that this abstraction is better because it's been vetted, but it may just be better because it's so easy to communicate it to a large number of people who are interested in economic growth. Even better: Sticking to these abstractions might provide some testable hypotheses.

Coase

The Coase theorem and its discontents seem central to both Eliezer's and Robin's AI worldviews. I'll set aside that topic for now, but I suspect that shoehorning the debates over the risks of unfriendly AI into a debate about Coase would be useful to economists: It's a common language for talking about win-win vs. win-lose.

Vetting

Eliezer hypothesized that economic growth models built around current human brains just wouldn't work for thinking about sped-up human-style brains. I hope I've demonstrated that these off-the-shelf models actually have some valuable insights. These might include:

1. The demand side can be a barrier to a *big* FOOM.
2. Other “weak links” on the supply side can be a barrier to *any* FOOM.
3. The fact that 2/3 of income reliably goes to workers even though most of today's output seems to be made by machines suggests that there's something on the supply side that makes humans strangely valuable.²
4. If just “speeding up the damn thing” is the same as an increase in w in Chad Jones's semi-endogenous growth model, then AI will have a big, transitory effect on the growth of technology, but *no impact on the steady-state growth rate*. So just how AI speeds up research matters a lot.
5. Computer scientists may have good ideas about the likely effects of AI on the semi-endogenous growth model's technology production function: Will AI be more like a rise in w , in a , in b , or in n ? Which kinds of near-term discoveries would cause us to rationally update our forecasts of AI-induced changes in w , a , b , or n ?
6. Transhumanists pushing for friendly AI may have normative preferences over the changes in w , a , b , or n . After translating the debates over AI into the language of semi-endogenous growth theory, they might want to encourage research in particular directions. For instance, based on my reading of the literature, the stronger we can make the “shoulders of giants” effect, the better off we would be from an economic perspective, setting aside political difficulties (N.B.: The debate over b was a major battle in the economic growth literature, for good reason).

I hope these comments are of some use to Eliezer and Robin. I certainly enjoyed spending a few days diving into their minds.

² Economists have hand-waved their way past this fact—but if economists had to take a position on the economics of AI (a la Robin's Robot Economy model) they'd have to take a stand on why human workers get 2/3 of output. This is just one example of how growth economists would be better off if they thought more about AI.