FROM: Ted Aronson

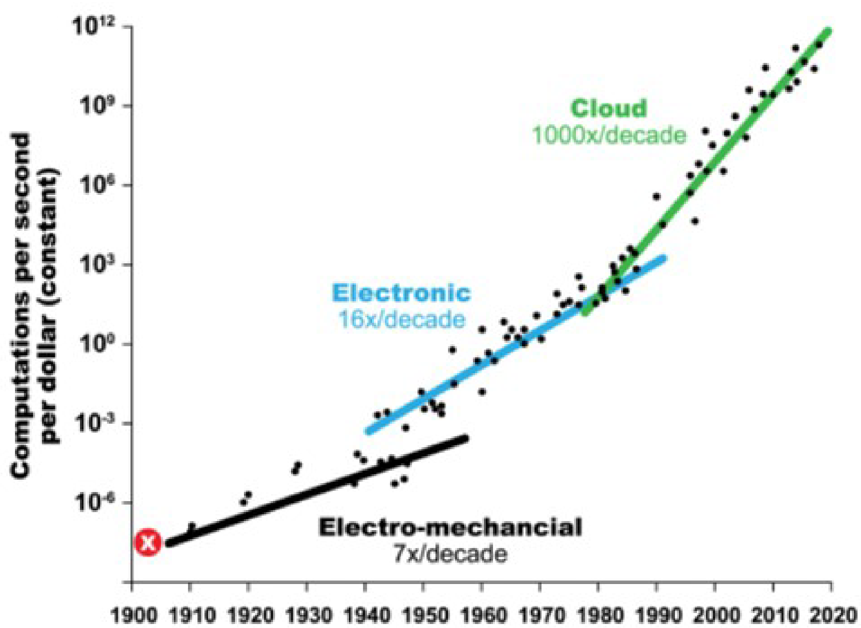
DATE: April 12, 2023

RE: **MARK MILLS**

If you liked Larry Siegel's review* of Mark Mills' latest book, *The Cloud Revolution*, you are going to like this piece!

Hint of what's to come:

EXHIBIT 2
QUANTITY OF COMPUTATIONAL SERVICES THAT CAN BE PURCHASED PER DOLLAR, 1900–2020 (LOG SCALE, BASE 10)



I strongly suggest you open the links within the article.

TRA
taronson@ajovista.com

P.S. Larry's *Fewer, Richer, Greener* is wonderfully optimistic about our future. Mark is of the same school — you will finish his comments feeling good!

*Holler if you'd like last year's review.

Conference Roundup: Mark Mills on the Cloud, the Robot Revolution, and Machines That Think

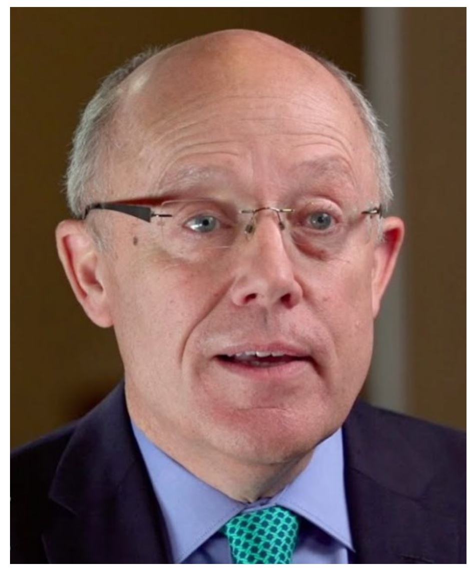
Edited by Laurence B. Siegel[1]
April 2023
Special to AJO Vista

On February 25, 2023, I heard Mark Mills, a physicist, venture capitalist, and author of *The Cloud Revolution*, which I reviewed in this space on April 6, 2022, speak at a meeting of the Global Interdependence Center in La Jolla, California. Without further ado…

Of tabulators and tractors

Mark Mills: Thank you, and thanks to the Global Interdependence Center.

The putative title of my talk is "The Technology of Money" because most of this conference so far has been about crypto, but I don't want to talk about the machines that facilitate finance — that facilitate transactions — that can trade wealth. Those machines are *tabulators*. I contrast these with *tractors*, which I'll get to in a moment.



Mark P. Mills
Source: City Journal

In 1952, for the first time in human history, a machine tabulated an election result. Walter Cronkite hesitated to release the results because the machine predicted that Eisenhower would win. He thought all the experts were right that Adlai Stevenson would win. He was wrong, the machine was right.

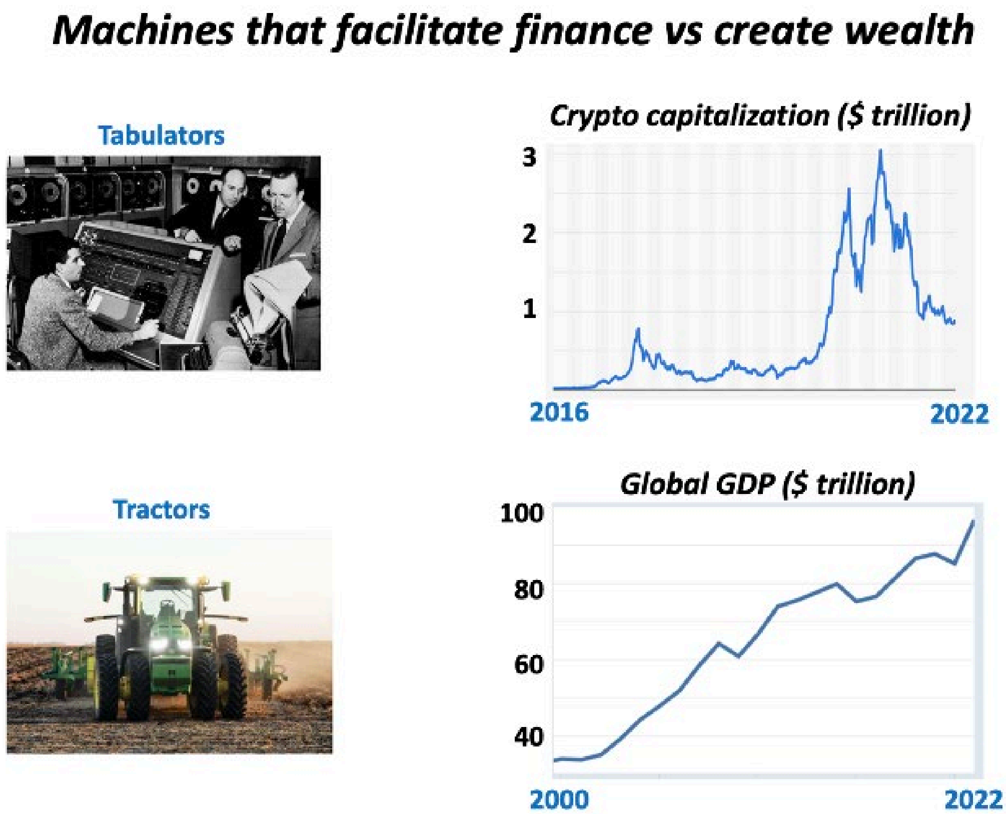
The very first commercial computer was bought by Bank of America to process checks. Although it was ridiculously expensive, it could process 40,000 checks an hour while humans were doing it at 200 checks per hour. So, it's beyond obvious why they used a computer to facilitate the recording of transactions moving through their system.

Cryptocurrencies are like that. They use a set of rules to process a transaction. Exhibit 1 is a graph of the total equity market cap, or value, of all cryptocurrencies combined from 2016 to date. The graph is about two weeks old and shows that total crypto market cap is around $1

[1] All exhibits are copyright © 2023 by Mark P. Mills, except for Exhibit 5 which was graphed by Laurence B. Siegel using FRED data, and Exhibit 6 which was constructed by Siegel and Stephen C. Sexauer. This article is a heavily edited transcript of a speech by Mills.

trillion. A trillion dollars is a lot of money, but the world's a big place, and crypto represents only 0.05% of the stored value of all forms of financial assets in the world.

Exhibit 1
"Tabulators" and "tractors" compared[2]



What I want to talk to you about, what I'm far more interested in, is the class of machines that includes tractors. Those that *create* wealth. Tractors are a classic example of a tool for wealth creation compared to what they replaced.

So, the really interesting issue for me, which I wrote about in my book, *The Cloud Revolution*, is: what are the underlying technology trends that facilitate not just exchange but also wealth creation? The world produces $100 trillion worth of goods and services every year — that is, global GDP is $100 trillion, and it's been growing. Its growth does not look anything like the cryptocurrency bubble. It's slow and steady and confers over time huge gains in standards of living for so many.

So let me focus on machines that increase productivity, in effect the "tractors" of the future. Most people know, if they studied economics, that productivity growth is what drives economic growth. Robert Solow received the Nobel Prize for documenting that fact. By

---

[2] Capitalization is a stock and GDP is a flow, so they cannot be compared directly.

"technology" he didn't mean Bitcoin, or ultrafast computers, or spacecraft. He meant technology in the broad, correct sense of the word: inventions and innovations that help you do more with less. Technology is the primary driver of the wealth and the productivity of the world.

There are patterns to technology's progress. My book is really about these patterns, but because I have limited time, I will illustrate them with just four examples.

The benefits of technological change come slowly

The patterns of technological innovations are non-linear, and take longer than most imagine. After radical innovations or discoveries emerge they then, always, take a while to be perfected. Of course, when they're perfected, there's a tipping point, a take-off point where they start to get used a lot. This is obvious, but what's interesting is that the timing of the patterns is pretty consistent over the ages. Introduction in 1908 of the Ford Model T marked the beginning of the auto age in earnest, but that was over 20 years after the first car was invented.[3]

Aviation grew from a concept to a big business in steps as well. The first scheduled airline passenger service in the U.S., operated by Western Air Express, took off in 1926 — again, about 25 years after the Wright Brothers. It took another 25 years, plus a few, to enter the jet age, which made flying fast and convenient — the Boeing 707 was introduced in 1957, and the sky was the limit after that. Computers — it's the same thing. Commercial computer mainframes came on the market roughly 20 years after the first electronic computer, and it was another 25 years before what we think of as the computer age took off.

The era we're in now, the age of the cloud, is just now getting moving, 25 or 30 years after the first cloud data center, Santa Clara Exodus, was built in 1995. This rhythm seems to apply to robotics too, and most other foundational technologies.

So, there's a predictable and understandable lag between the development of a new technology and its widespread adoption and influence on the economy. Solow famously said in 1987, "You can see the computer age everywhere but in the productivity statistics." Twenty years later you *could* see it in the productivity statistics.

Higgledy piggledy growth[4]

So let me tell you where we are today with computing power in economic terms. Any economist wants to know the progress in how much of a good or service you can buy for a

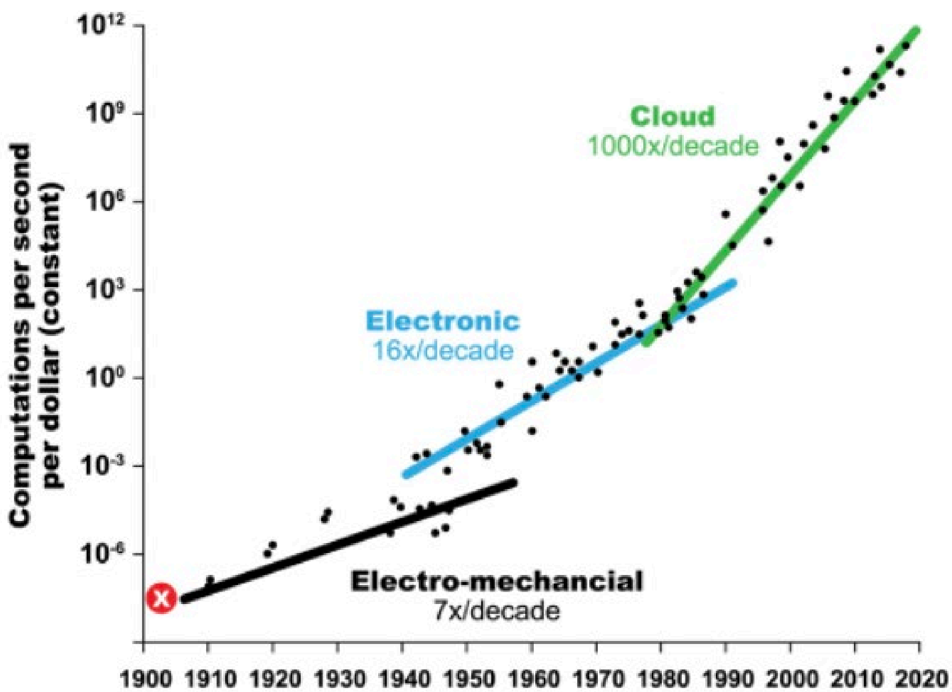
[3] Karl Benz, in 1886, is usually credited as the foundational inventor, although France's Nicholas Joseph Cugnot built a steam-powered tractor in 1769. See "Who Invented the Automobile?" (undated), Library of Congress, https://www.loc.gov/everyday-mysteries/motor-vehicles-aeronautics-astronautics/item/who-invented-the-automobile/. A century earlier, in the 1670s, Ferdinand Verbiest built a small prototype (too small for a driver and unable to be steered) of a self-propelled steam vehicle.

[4] The oldest of my readers will know exactly what this refers to. Younger folks can Google it, or ask ChatGPT if you want a much longer and probably inaccurate answer.

dollar. In the case of computing power, what you're buying is computations per second. This is true whether you're buying the machine, leasing it, or using it remotely in the cloud.

The history of modern computation is one of growth at increasing rates, as Exhibit 2 shows. It's an exponential curve. During the first age of electromechanical computing, the economic value of these machines increased sevenfold per decade. You got seven times more value, measured as the number of computations per second, each decade. Later, the first electronic age increased value by 16-fold per decade, a much faster growth rate.

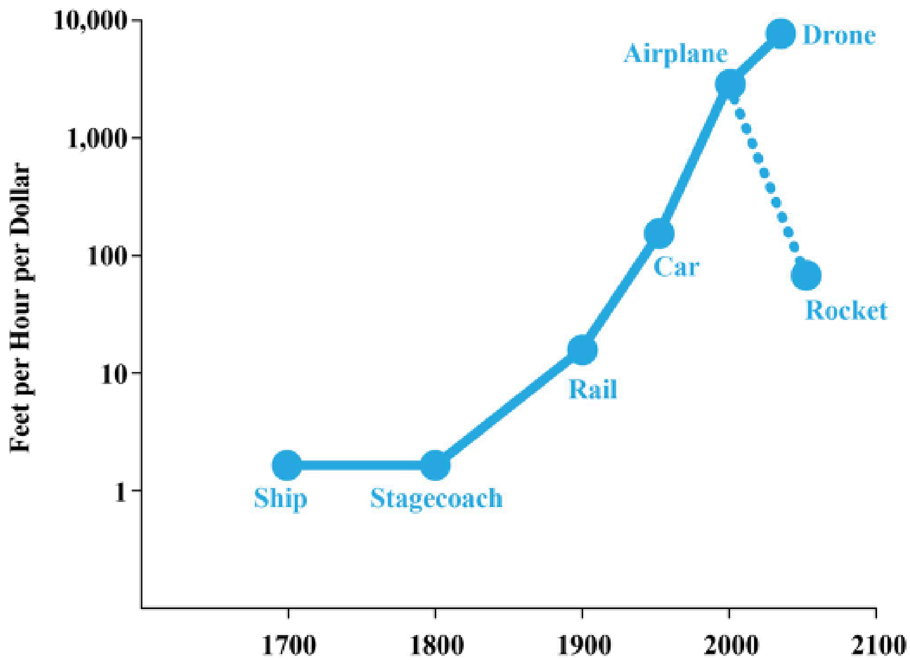
EXHIBIT 2
QUANTITY OF COMPUTATIONAL SERVICES THAT CAN BE PURCHASED PER DOLLAR, 1900–2020 (LOG SCALE, BASE 10)



But, in the age of the cloud, value has been going up 1000-fold per decade.

You can compare this to the improvement in the economic cost or value of transportation services — how much it costs to move goods or to go somewhere. It's also an exponential curve. Exhibit 3 shows that the economic efficiency of transportation services also grew by a thousandfold, but that happened over the entire 20th century — it took 100 years. We're now doing that every 10 years in computation services. This is consequential in ways that are hard to sort out.

EXHIBIT 3
AMOUNT OF TRANSPORTATION THAT CAN BE PURCHASED PER DOLLAR, 1700–2100 (PROJECTED), LOG SCALE (BASE 10)



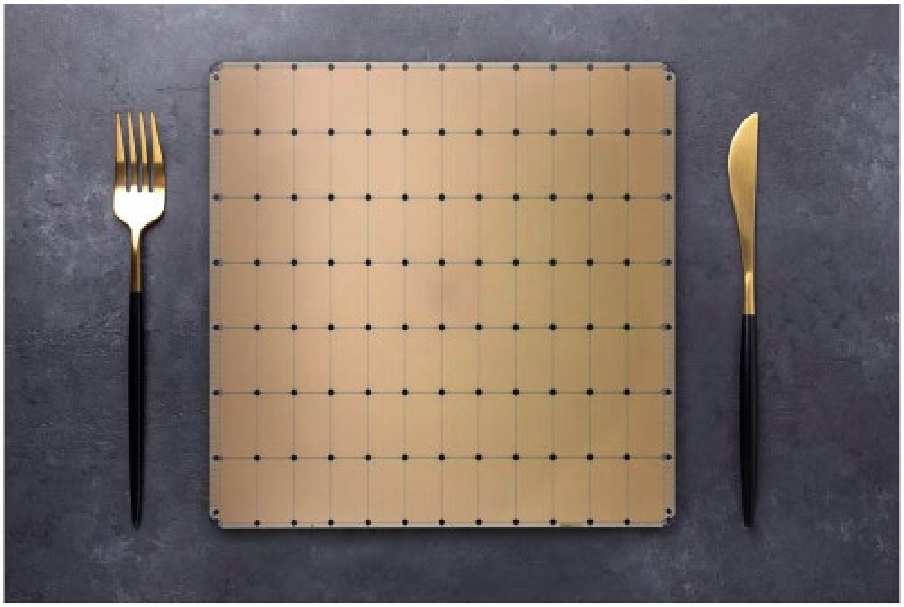
AN INFERENCE CHIP THE SIZE OF A DINNER PLATE

Coming back to the machine realm, let's consider a little-known company called Cerebras. It's a commercially viable post-startup company. It manufactures a new class of inference chip — not a computer chip *per se*. In the world of artificial intelligence, you're doing inference, not computation. I know roughly where I want to drive. The answer isn't computational in the sense of 1 + 1; it's continuous adaptation to changing circumstances. To do that, all the early machines used computation chips to perform inference but used brute force, which is very inefficient, like using a helicopter to fly across the Atlantic.

Or you can build an inference chip. Graphics processing units (GPUs), what NVIDIA began to produce 25 years ago, are inference chips or engines, which was the class of processing unit that enabled the efficient use of artificial intelligence software. We get things like Siri and Alexa because they do pattern recognition on your smartphones, which requires inference.

The inference chip in Exhibit 4, made by Cerebras, is the size of a dinner plate. It came into commercial being 25 years after the first AI chip, echoing the pattern I discussed earlier. This chip processes images, and does inference 10,000 times faster than the fastest NVIDIA GPU. It's now a commercial product, not an idea.

Exhibit 4
Cerebras WSE-2 computer chip, shown in spatial context



For those of you who, like me, also work in the energy domain, you'd be interested to know that this one chip uses 15 kilowatts of electricity. This is the capacity of three typical houses, or one very large house in La Jolla, running 24/7.

This chip has 3 trillion transistors on it. It's an astonishing engineering feat. It allows you to do inference faster, easier, and cheaper even though it's very expensive, each costing about $1 million.[5] A group of these chips was assembled in December 2022 — 16 of them — to make what's called an exaflop supercomputer. (An exaflop is one quintillion, or 10^{18}, floating-point operations, or flops, per second.) The previous exaflop supercomputer was a $600 million, 47,000-processor machine at the Oak Ridge National Labs that requires 25 megawatts to run. The Cerebras exaflop supercomputer using 16 of these dinner-plate chips requires just 2 megawatts and costs $30 million to build. They built it in one week.

Where is this supercomputer going to reside? In the cloud. It's not going to sit in your smartphone. It will sit in the cloud and it will do inference. It will make ChatGPT smart. ChatGPT is already smarter than Siri and Alexa by a lot. All of us had the experience of trying to convince Alexa or Siri to understand English. ChatGPT is better. ChatGPT is still C–, maybe D+.[6] This additional leap in AI processing power will make chatbots, which are central to

---

[5] This is one three-millionth of a dollar, or one thirty-thousandth of a penny, per transistor.

[6] At the time Mills said this, ChatGPT-4 had not yet been released. ChatGPT-4 is starting to approach the B– output of an 11th grader, which is pretty good for a machine. It's still what Mills calls "really good automation," not intelligence. Machine learning is a better moniker: it's a machine, and it does appear to be learning. Let's not get over our skis.

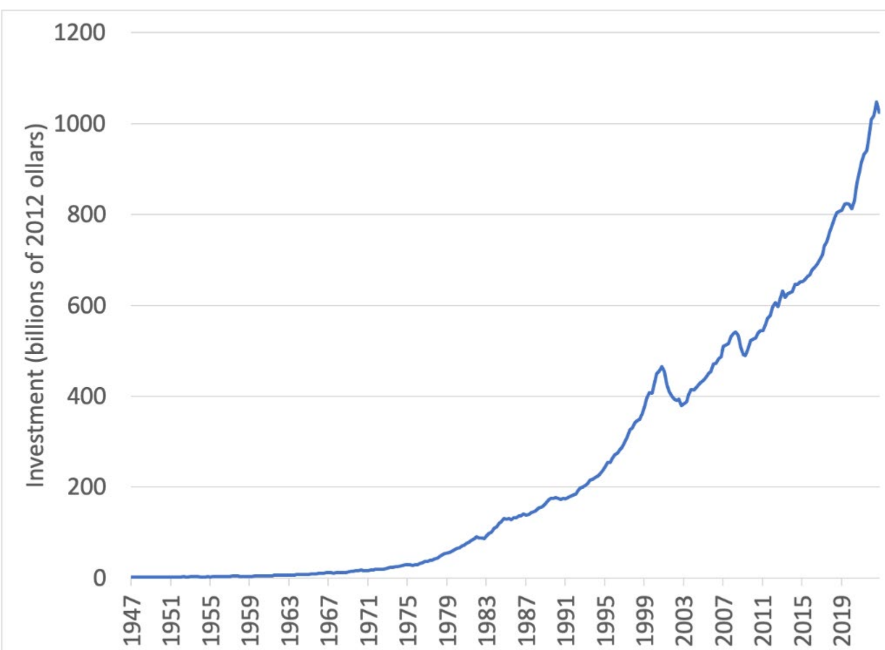
making a lot of financial transactions easier, actually useful, and a lot else besides. I think this is consequential, if not revolutionary.

THE CLOUD AS AN ENGINE OF ECONOMIC GROWTH

The essential contention of my book is that we are at a dislocation, a pivot in history in the functionality, reach, scale of the cloud that is humanity's newest infrastructure. The cloud is not the internet — the cloud uses the internet. The cloud isn't about doing computation (though it can) but providing inference and advice; it isn't about just storing data or cat videos, though it does that too. Of course the cloud stores the unbelievably massive amount of data produced every day, about so many aspects of commerce. The cloud is a utility for providing advice and support to the economy at large. Of course it is the engine that makes crypto work, but that's just one use of it, and not a very large one.

Exhibit 5 shows spending, in constant (2012) dollars, by U.S. businesses on IT services and hardware starting in 1947. There's no sign that this is peaking soon or even leveling off.[7] Businesses don't buy non-useful things. Businesses are spending a trillion a year now, in the United States alone, to buy information processing and the means of acquiring information. That's meaningful.

EXHIBIT 5
PRIVATE FIXED INVESTMENT IN INFORMATION PROCESSING EQUIPMENT AND SOFTWARE, 1947–3Q2022



[7] However, the growth rate has decreased, as you'd see if the y-axis used a log scale. (I used an arithmetic scale to emphasize the sheer volume of investment.) But growth rates invariably slow as the denominator becomes enormous, 'hiding' the continuing rapid rate in absolute growth.

Capital spending on global cloud infrastructure now exceeds global capital spending by all electric utilities combined, even though we're electrifying everything. And the growth rate of spending on the cloud is far greater than the growth rate of spending on electric utilities. If you think in real estate terms, the cloud is a network of buildings and communication systems and end-devices, and the buildings at the core of the cloud, the datacenters, are the size of skyscrapers in square footage — a million square feet under one roof. They cost about the same per square foot to build as a skyscraper. They rent for five times as much per square foot as a skyscraper, and they consume 100 times as much power per square foot as a skyscraper. The world has built about 1,500 skyscrapers over the past century, and over 5,000 enterprise, skyscraper-class, datacenters in the past couple of decades; the latter is expanding far faster than the former. It gives you some sense of where the world's cloud infrastructure is going, and the scale of its energy appetite.

Much more important is the economic impact of the marginal dollar spent on the cloud versus spending on other things. The marginal dollar spent on steel, a new car, any new physical asset typically gives you a new asset that is better, but only a little bit better. But each marginal dollar spent on the next unit of IT hardware or service is not just a little bit better. It's getting better at the rate I showed you earlier — a thousandfold per decade. It's an astonishing economic accelerator.

That will, inevitably, show up in the macroeconomic data. The cloud is blowing away the traditional production function.

Not enough people? Build robots...lots and lots of them

Let's turn to a couple of ideas on labor and employment that relate to our previous speaker, Patrick Harker [president of the Philadelphia Fed]. He said we don't have enough people, and I agree with that. In fact, Elon Musk has said the same. The world is looking at a very bad demographic path, for lots of reasons, as couples are choosing to have fewer and fewer children, in many cases none at all.

One of the functions of the cloud — one of the functions of AI — is to address the longstanding challenges in labor markets. AI deployed in robots is already starting to migrate into these markets, especially those that involve manual labor. The absence of robots in businesses is startling. Some 90% of the manufacturing businesses in the United States don't have a single industrial robot. You can invert that: 90% of the industrial robots are in 10% of the businesses. This may be because robots, until now, have only been good at doing highly repetitive tasks that very large businesses undertake.

Until now, there haven't been robots that are useful, mobile, and adaptive, and can work alongside people in multi-function tasks that characterize the majority of industrial businesses. The class of mobile robots is now at a take-off point, again coming about 25 years after the first viable ones were built. Nearly one million such units shipped last year.

The robot of the last era was the robot that you bolted down and set to doing some kind of highly repetitive task. The very first robot to do that was in 1961. It was called Unimate. In fact, that year it showed up on Johnny Carson's TV show and was also installed for the first time in a General Motors factory, making that factory the most productive one in the world. Every automaker on the planet chased that. The vast majority of industrial robots installed thus far have all been focused on and effective at doing the highly repetitive mass-production tasks with the automotive industry the utterly dominant (stationary) robot buyer.

In a coincidence, one of the first commercially available useful mobile robots showed up 61 years later *on the same stage* that Unimate appeared on, with Johnny Carson's stage now being used for the Jimmy Fallon show. You can see "Spot," a walking dog-like robot, appear with Fallon here.[8] Other examples of current robotic advances are here,[9] showing an industrial use case for Spot — safety inspections — and here,[10] an example of a pre-commercial anthropomorphic warehouse machine from Agility Robotics.

But if you really want to be impressed, check out the motor skills of Boston Dynamics' Atlas robot. You can see Atlas do some funky fifties-style dancing with some "friends", navigating a parkour course, and being helpful around a workshop.[11] "His" feats in the last video seem to involve reasoning as well as athleticism. Atlas, and its competitors, are not imaginary concepts but machines on the cusp of commercial viability with profound implications for productivity in markets where there is an endemic and worsening labor shortage.

How do we get all these jobs done?

Spot and Atlas, plus ChatGPT, and all the dozens of similar machines in pre-commercial development, are in fact revolutionary. Five years from now, the economists at the Federal Reserve will look back on this time period as a pivot in which we solved an interesting problem. How do we get the jobs done that are underserved, undesirable?

Exhibit 6 uses Federal Reserve data to categorize all of U.S. labor into a two-by-two matrix: routine versus non-routine, and cognitive work versus manual work.

---

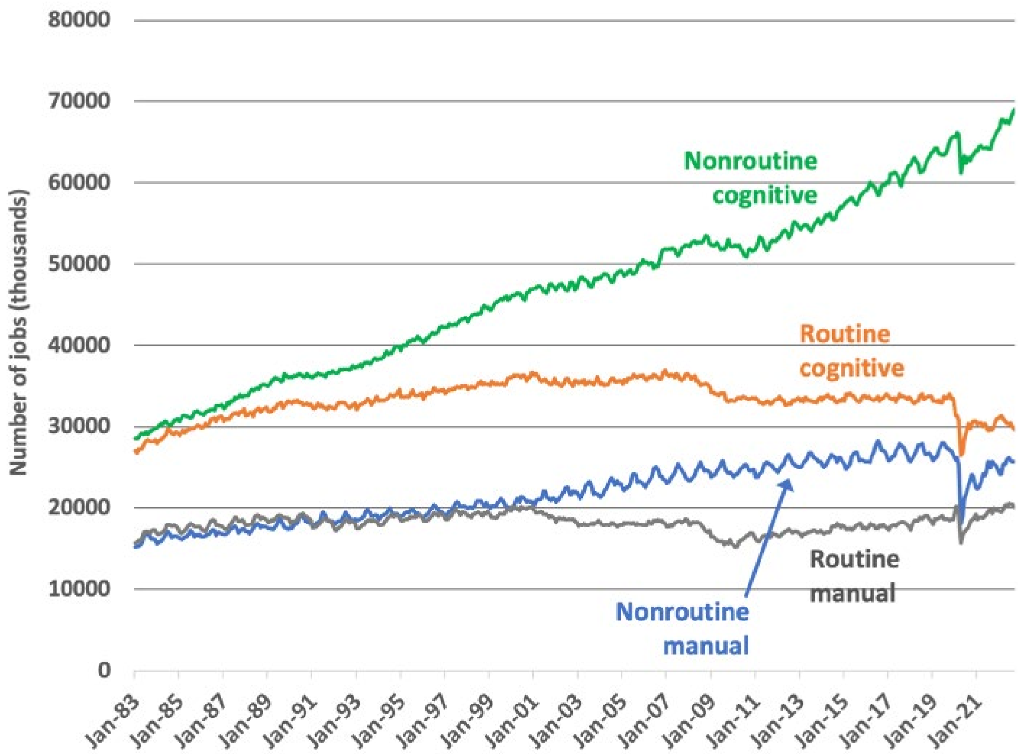
8 https://www.facebook.com/watch/?v=345579204239794

9 https://www.youtube.com/watch?v=i1iPOKoI2gE

10 https://www.youtube.com/watch?v=hD4NiXg861c

11 https://www.youtube.com/watch?v=fn3KWM1kuAw, https://www.youtube.com/watch?v=tF4DML7FIWk, https://www.youtube.com/watch?v=-e1_QhJ1EhQ&feature=youtu.be

EXHIBIT 6
EVOLUTION OF JOB QUALITY IN THE UNITED STATES, JANUARY 1983–AUGUST 2022



Source: Siegel and Sexauer (2023), using data from FRED, Federal Reserve Bank of Saint Louis.[12]

You can see that, over the last 40 years, there was essentially zero change in the absolute number of people engaged in routine tasks in both the manual and cognitive labor categories, despite a big increase in population. All the growth in labor has been in non-routine work, both cognitive and manual. If we don't have enough people to do the non-routine jobs where demand is growing, we can find them by moving them out of routine task. This latter transition is made possible by amplifying routine labor efficacy, i.e., reducing the number of people doing routine work (both manual and cognitive) and adding more robots, both physical machines and virtual AI robots like ChatGPT. Routine tasks are precisely where those technologies are best suited.

Doing that frees up capital, frees up people from doing the worst jobs, enhances productivity, and allows wage growth without wage inflation. So I think the Fed's challenge in tamping down inflation will be solved — I'll be slightly provocative–much more effectively by facilitating the deployment of new machines than by manipulating interest rates and raising business taxes.

[12] Siegel, Laurence B., and Stephen C. Sexauer. 2023. "Longer, Healthier, Happier: Why Working Longer Improves Almost Everything." Forthcoming in *The Journal of Retirement*. Comment draft here.

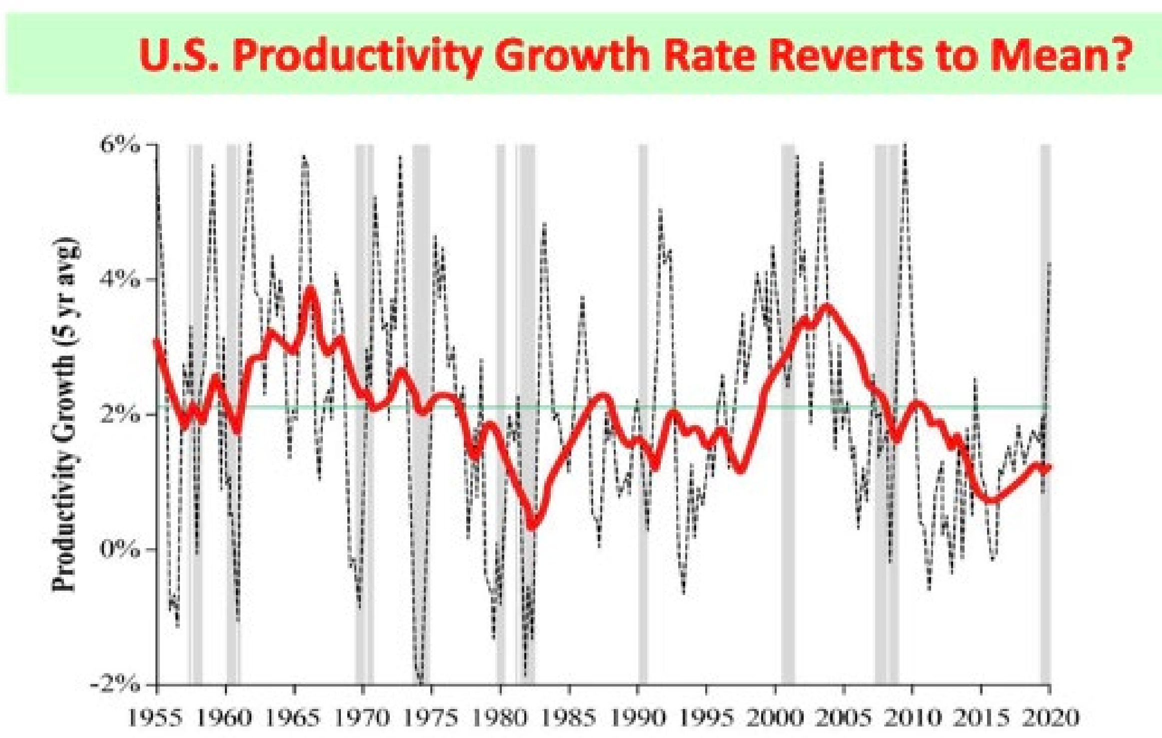
MACRO OBSERVATIONS AND CONCLUSION

Two quick macro observations: The job openings ratio — the ratio of U.S. jobs available to the number of people that are available to be employed — just flipped. We now have more jobs than people available. You all know that. And, second, the demographic shape of the global labor pool has flipped: the working-age population is shrinking as a percentage of the total population. This is happening in just about all countries, but is most visible in the United States, Europe, and China.

We're going to need ChatGPT-class machines on steroids, and we're going to need many more robots. We will get them; dozens of innovators are building them and even the most famous modern machine innovator, Elon Musk, has joined the pursuit recently announcing Tesla's plan to build an autonomous, mobile, humanoid robot. The announcement is worth seeing.[13]

All this means that the average productivity growth rate will tilt back up. Exhibit 7 shows the five-year trailing productivity growth rate in the United States. The underlying technology trends tell us that it will rise to the historical average and beyond. The data are beginning to show that already.

EXHIBIT 7
U.S. PRODUCTIVITY GROWTH RATE, TRAILING FIVE-YEAR AVERAGE, 1955–2020



[13] https://www.youtube.com/watch?v=UXHoWNfjJYM

So, I expect these technologies will bring profound dislocations, and wealth creation. There's more pent-up growth than most people realize. Economists, not to be unkind because most of you are economists, have historically been profoundly bad in predicting economic dislocations caused by technology. But they've been very good at hindcasting the impacts of technology dislocations once they happen and framing the resulting economic rules of the road. Let's not insult them, because those rules do matter. Thanks.

Mark P. Mills *(speaker) is a physicist, a Manhattan Institute senior fellow, a faculty fellow at Northwestern University, and partner in Montrose Lane, an energy-tech venture fund. He is author of The Cloud Revolution (2021), Digital Cathedrals (2020), and Work in the Age of Robots (2018), and co-author of The Bottomless Well (2006). He served as chairman and CTO of ICx Technologies, and co-authored a tech investment newsletter, the Huber-Mills Digital Power Report. Prior to that, he served in the Reagan White House Science Office and worked in the commercial nuclear industry. He began his career as an experimental physicist and development engineer in microprocessors and fiber optics, earning several patents while working at Bell Northern Research and RCA. He holds a BSc, Honours, in physics from Queen's University, Canada. His website is https://www.tech-pundit.com.*

Laurence B. Siegel *(editor) is the Gary P. Brinson Director of Research at the CFA Institute Research Foundation, economist and futurist at Vintage Quants LLC, and the author of* Fewer, Richer, Greener: Prospects for Humanity in an Age of Abundance *and* Unknown Knowns: On Economics, Investing, Progress, and Folly*. He works as a writer and independent consultant and may be reached at lbsiegel@uchicago.edu. His website is http://www.larrysiegel.org.*