Winter Heating

Ubiqus
61 Broadway – Suite 1400 – New York, NY 10006
Phone: 212-346-6666 ◆ Fax: 888-412-3655

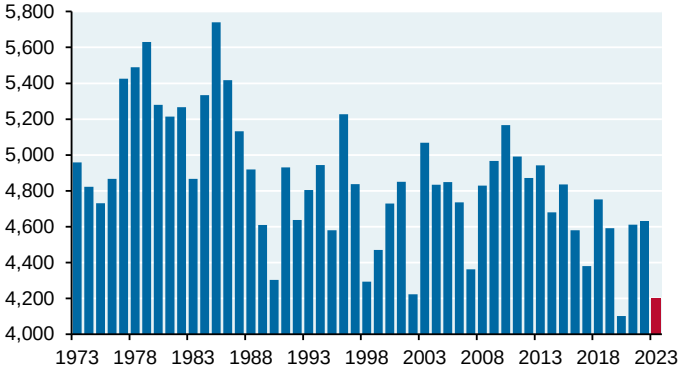
Winter Heating

[START RECORDING]

FEMALE VOICE: This podcast has been prepared exclusively for institutional wholesale professional clients and qualified investors only as defined by local laws and regulations. Please read other important information, which can be found on the link at the end of the podcast episode.

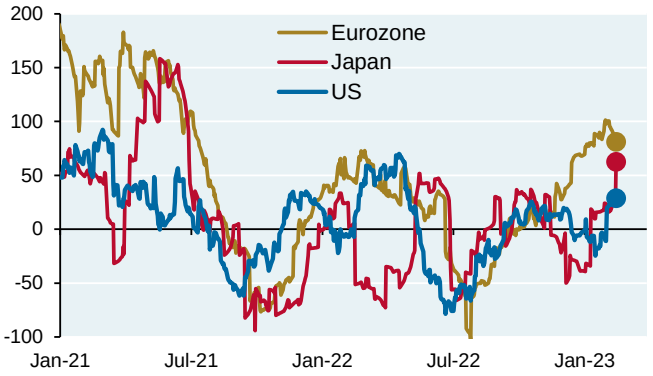
MR. MICHAEL CEMBALEST: Welcome to the February 2023 Eye on the Market podcast. This one is called Winter Heating. Most of what we're going to be talking about are the battles between the new large language models, which you're reading about everywhere, a quick few comments on the markets and the Fed. One of the warmest northern hemisphere winters in decades has coincided with a flurry of positive economic surprises in the U.S., Europe, and Japan. The U.S. list of positive surprises is a long line, from employment, retail sales, manufacturing, jobless claims, a miniscule high-yield default rate, capital spending projections are stable, et cetera, so definitely a surprise in terms of how strong the U.S. economy is in January and February.

Historical heating degree days in winter for the Northern Hemisphere, Population-weighted days, Jan-Feb



Source: J.P. Morgan Global Commodities Research. February 16, 2023.

Positive economic surprises to start the year in the OECD
Economic surprise index

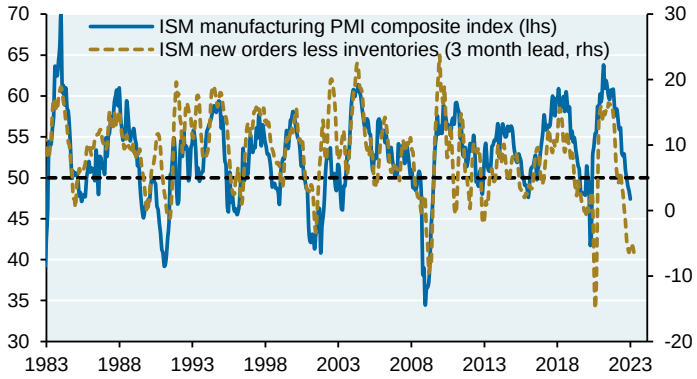


Source: Bloomberg, JPMAM. February 16, 2023.

The problem is that policy rates aren't normalized yet, and looks like we may peak closer to 5 and a quarter to 550 on the funds rate before the Fed pauses. So we still see some weakness ahead in our preferred leading indicator, which is looking at new orders versus inventories. And the bottom line is it looks like we've got two to three more Fed hikes ahead and a mild recession whose likelihood and possible

severity is probably shrinking.

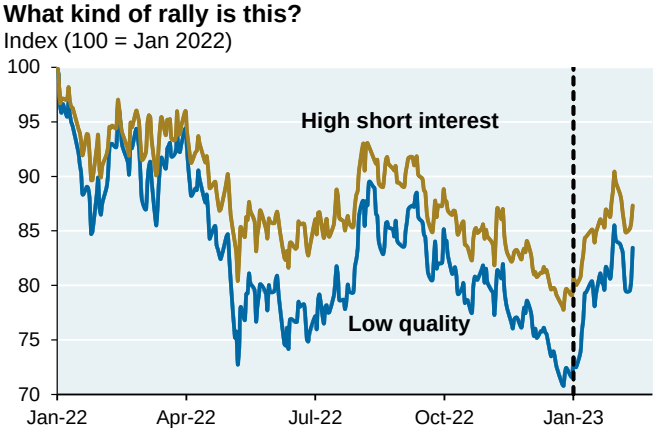
US new orders less inventories predicting ISM downturn



Source: Bloomberg, JPMAM. January 2023.

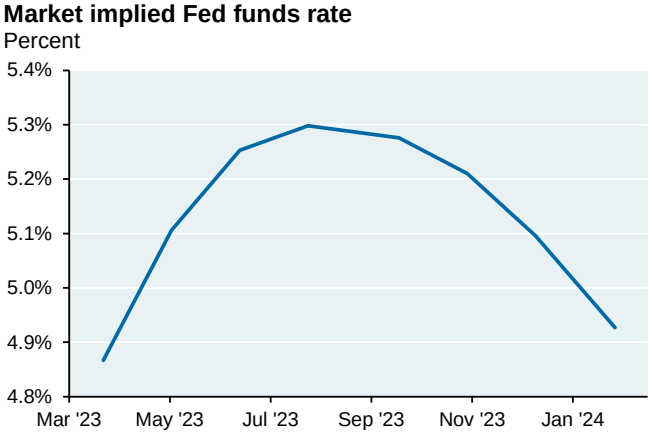
What makes us not really want to chase the equity market rally here is that the rally so far this year has been based, at least in part, on appreciation of some low-quality and high-short-interest stocks, a huge explosion in the money supply both in Asia, and the view that the Fed hikes aren't going to do that much damage to the economy, which I think is still flowing through.

What kind of rally is this?



Source: Bloomberg, JPMAM. February 16, 2023.

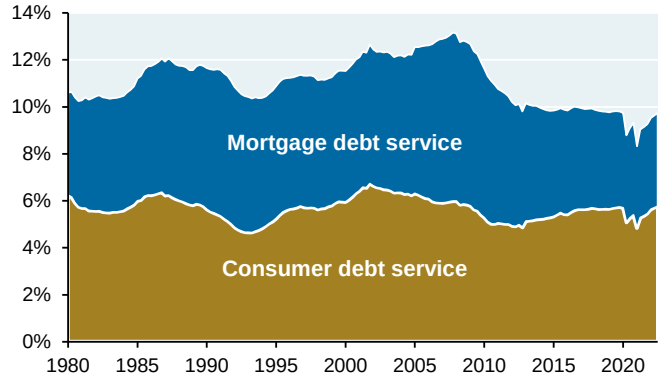
Market implied Fed funds rate



Source: Bloomberg, JPMAM. February 20, 2023.

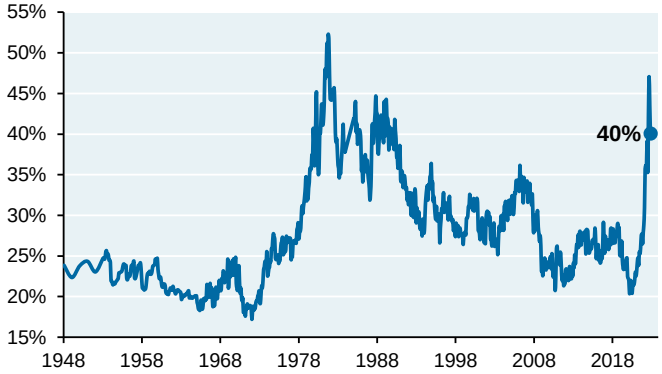
We did not have the view there was going to be a severe recession this year; we thought it would be a mild one. And we still think that's the case. And as we're looking at leading indicators for earnings, it does look like there is some downside for both earnings and margins in the month ahead, but nothing too catastrophic.

Household debt service share of disposable income
Percent



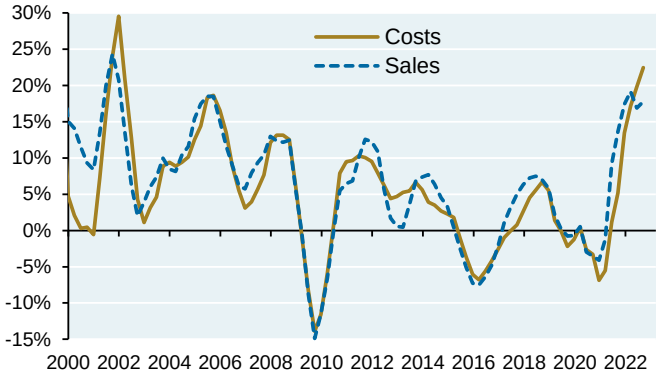
Source: Federal Reserve Bank, JPMAM. July 2022.

Annual mortgage cost for new homebuyers
Percent of household income



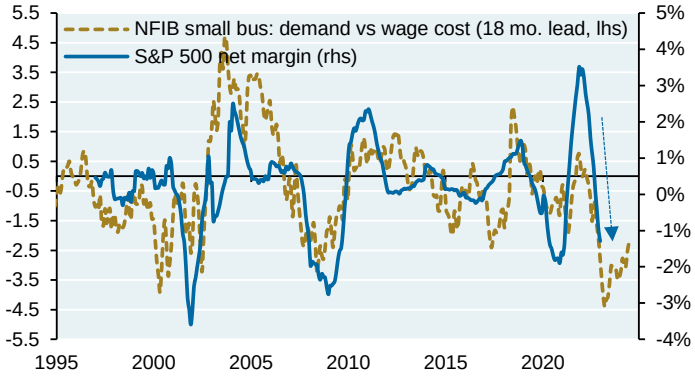
Source: Census, Shiller, Freddie Mac, Bloomberg, JPMAM. December 2022.

Costs rising faster than sales
Percent, y/y



Source: Factset, JPMAM. September 30, 2022. Universe: CRSP all cap.

Leading margin indicator predicts downside ahead
Index composite Percent, y/y change



Source: Morgan Stanley Equity Research. January 2023.

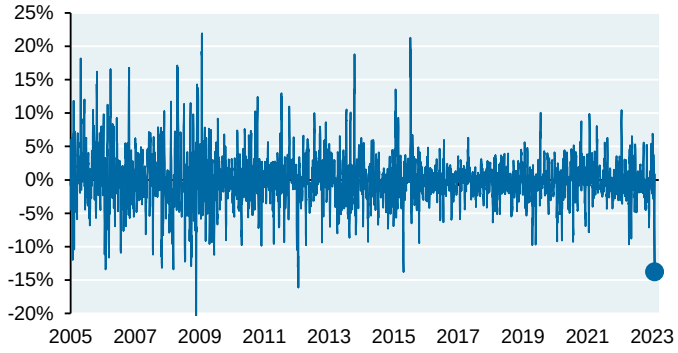
The centerpiece of today's podcast and our Eye on the Market this week is on these large language models, which have been impossible to ignore over the last couple months. I was at our annual client investment conference in Miami a couple weeks ago was listening to Sam Altman from OpenAI talk about ChatGPT. And it was on the same day that Google had rolled out Bard, its own large language model, where there was a perception of a botched rollout by Google.

And it resulted in one of the largest weeks of underperformance of Google versus Microsoft in a decade and one of the largest since Google's IPO in 2004, which was ironic because just last month, one of Google's large language models passed the U.S. medical licensing exam for the first time. It was reportedly the first one of these

models to be able to do that.

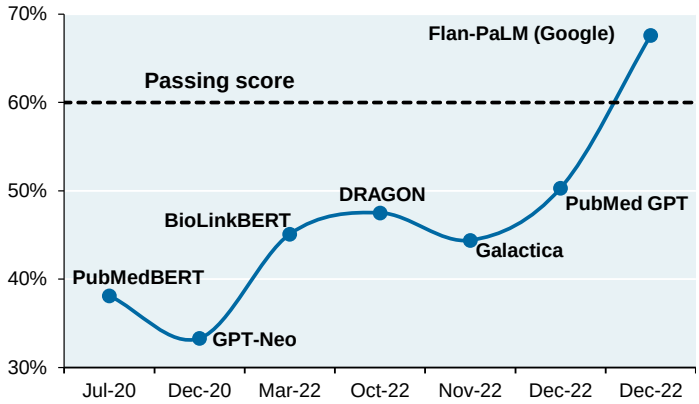
Google's botched Bard rollout resulted in one of its largest weeks of underperformance vs Microsoft on record

Percent, Google - Microsoft weekly total return



Source: Bloomberg, JPMAM. February 17, 2023.

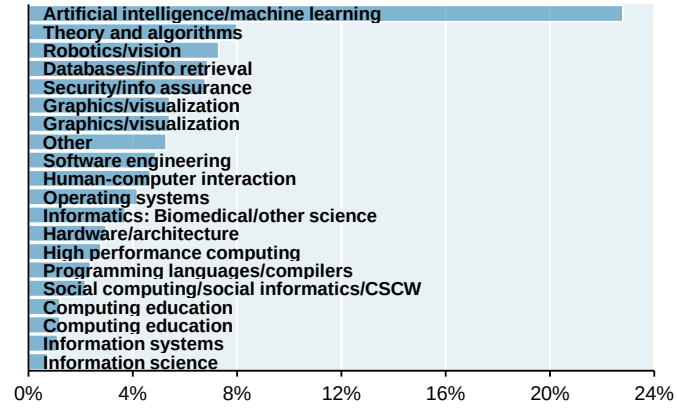
US medical licensing exam



Source: Technical University of Denmark. December 20, 2022.

So I want to give you some big picture thoughts on all of this large language model stuff. Artificial intelligence is attracting a ton of VC money and mindshare among computer science Ph.Ds. I've been very critical of unprofitable innovation over the last couple of years, whether it was the metaverse or hydrogen or these buy now, pay later companies or crypto or anything. But I feel differently about this stuff without getting into the details of pre-IPO evaluations for specific companies. I think these models are going to result in greater productivity benefits and disruption than those other things, although the bar is admittedly low when you're talking about things like metaverse and crypto.

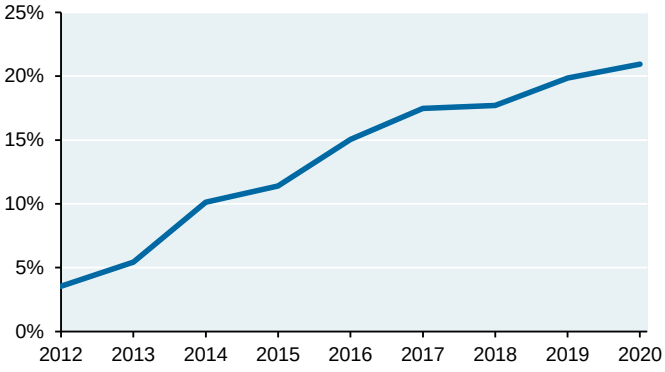
Share of US computer science PhDs by specialty



Source: AI Index Report, Stanford. 2021.

Share of total VC investment into Artificial Intelligence

Percent



Source: OECD AI Policy Observatory. 2021.

Large language models are essentially conventional wisdom machines. They don't know anything other than what has already been documented in digitized human experience. They are not artificial intelligence the way that you conceive of it in terms of it coming up with something new that hasn't been thought of yet.

That said, there's billions and billions of dollars of market cap in companies that traffic in the packaging and conveyance of conventional wisdom all the time. And so what is interesting is that artificial intelligence in these large language models aren't likely to kind of break any new ground intellectually, but they have the potential to displace a lot of the conventional thinking and conventional wisdom companies that exist right now.

So before we get too carried away, I want to be clear about a few things. These models make tons of mistakes, tons of them. And some of them are kind of hysterical. They recommend books that don't exist, they don't know what year it is, they believe certain countries have left the EU and it hasn't happened, they make up numbers in earnings reports, they write essays on the benefits of adding woodchips to your breakfast cereal. The list is endless, and some AI people describe these models as sarcastic parrots because all they're just doing is repeating things that they see, and sometimes not correctly.

And remember, it was just last November when Facebook/Meta rolled out Galactica, which was a large language model designed to help researchers summarize academic papers and solve math problems and write code. It was unable to distinguish truth from falsehood. It produced articles about the history of bears in space and got yanked after three days of intense criticism from the scientific community.

As another sign of how far there is to go here, Stack Overflow, which some of you may have heard of, is a question-and-answer site that a lot of developers and programmers use, I've even used it to help me with some Visual Basic code, they put a ban on ChatGPT submissions at that website because the rate of getting correct answers from it wasn't good enough.

And so we're still in a world where there's a lot of errors that come out of this stuff, and you can make it, you can

make these models produce garbage. Some researchers at Northwestern trained a model to write fake medical research abstracts, which both other models and humans couldn't figure out were fake.

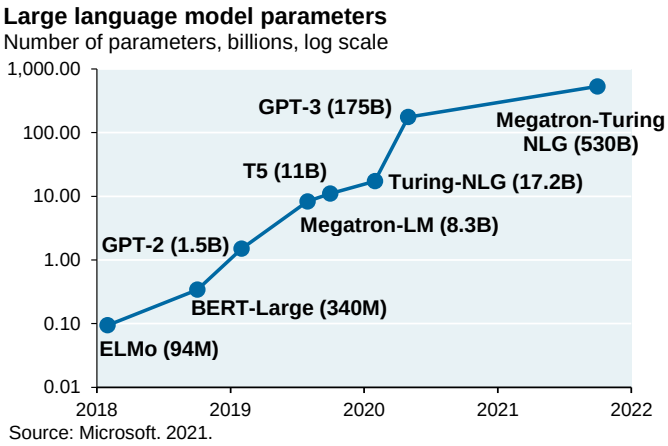
There's also a lot of hype in this space. One of the grandfathers of machine learning in 2016 advised hospitals to stop training radiologists because we wouldn't need them anymore since deep learning would be better. And here we are a few years later, and it turns out machine learning for radiology is a lot harder than it looks. And artificial intelligence in that kind of thing is best used when complementing doctors instead of replacing them. All of these errors and bizarre things that large language models do are referred to as hallucinations, and you can see why.

Now with that, these models are making progress on a lot of well-specified tasks. And despite what Stack Overflow to ChatGPT, there are a couple of new companies that are being rapidly embraced by developers. GitHub has something called a copilot tool, added 400,000 users in its first month, now has over a million users who use it to help them with 40% of their code. Tabnine is another artificial intelligence-powered coding assistant backed by a lot of the who's who in Silicon Valley, also have a million users. And Microsoft has an advantage here, 'cause it's got a partnership with both OpenAI and it owns GitHub.

And there have been just in the last few months some interesting practitioner and academic analyses of what you can do with these models. They outperform sell-side analysts when picking stocks. They show promise in putting together short-term, long-term, long/short trading strategies based on synthesizing comments from CFOs in conference call transcripts. They are already improving audit quality at the big accounting firms if you use the frequency of audit restatements as a proxy for that. The University of Florida created something called GatorTron to extract insights from terabytes and terabytes of clinical data to help them with medical research.

And in law firms, this is looking interesting. These models have correctly gone through court cases and predicted with the judgements were going to be. They have passed bar exams. They have begun to be used to draft contracts and conduct legal research, et cetera, et cetera. And Microsoft just

released MegaTron, which is the largest one of these things to date. So there is evidence that if you narrow down the scope of what you're asking these models to do, that they can be more productive and do a much better job.



Source: Microsoft. 2021.

One of the big questions that came out of this, as I mentioned earlier, Google's stock got pummeled after this rollout. What's going to happen to the profitability of the search business? Microsoft's CEO says the gross margins of these, of search is going to drop forever, and Sam Altman is referred to lethargic search monopolies. We know who he's talking about when he says things like that.

We'll see. I mean Google knows a lot about machine learning and artificial intelligence. They were the ones that designed some of these initial transformer programs. There's a lot of machine learning that already is going on within Google Search. And I anticipate a pretty robust response from them soon regarding their capabilities after what happened.

That said, the search economics do look more challenging. Google's operating margins are about 24%. And that includes YouTube, so obviously without that, they're higher. But anything that they offer with respect to large language models being integrated into the regular search offering would sit on top of their existing cost structure.

Now ChatGPT's cost structure is pretty high because they're not completely vertically integrated, and depending upon the number of words generated per query, the model size, and

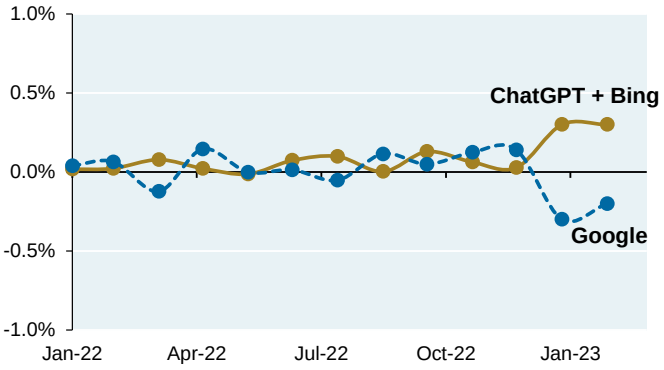
their computing costs, it can get very expensive for them. And some of the estimates we've seen is anywhere from five to ten times more expensive per query compared to the standard Google search query.

But I think it's important to remember, first, Google has announced that when and if they incorporate large language models into the search engine, it would be a lightweight version instead of its heavy-duty version. So I think the ChatGPT cost estimates probably overstate the impact on Google's operating margins.

Also, there are things called sparse models. Right now if you use ChatGPT 3, all 175 billion parameters are used to generate a response, and sparse models are built narrowing the field of knowledge that you need to answer a question, and they require a lot less training energy, a lot less computing energy, and actually can work faster.

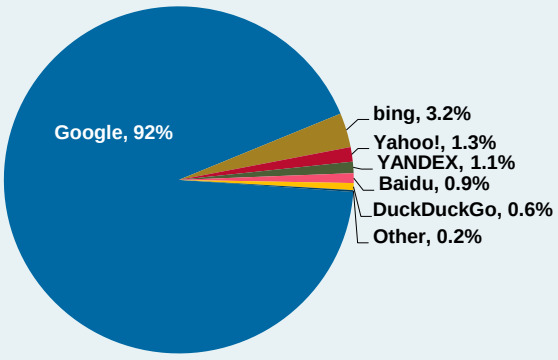
I know it's tempting to kind of look at the pace of innovation and see a big threat to Google here on the search side. So far, the ChatGPT/Bing combination has taken like .3% market share on search traffic from Google, whose base market share is above 92%. So I just think it's important to look at the numbers to keep these things in context.

Search and chat website traffic
Percent, m/m change



Source: Morgan Stanley Equity Research. 2023.

Search engine market share, Jan 2022-Jan 2023
Percent



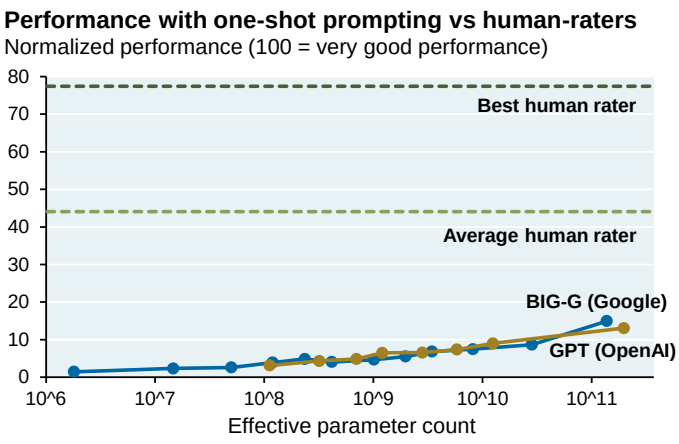
Source: statcounter Global Stats, 2023

What's the future of these things, these large language models? One of the more interesting projects that's underway is, despite all the hyperbole of what these things can and can't do, there is about 400 researchers from Google, OpenAI, and 100 other firms, and they've put together something called the BIG-bench, which refers to the big benchmark. And they've all created about 200 tasks for large language models

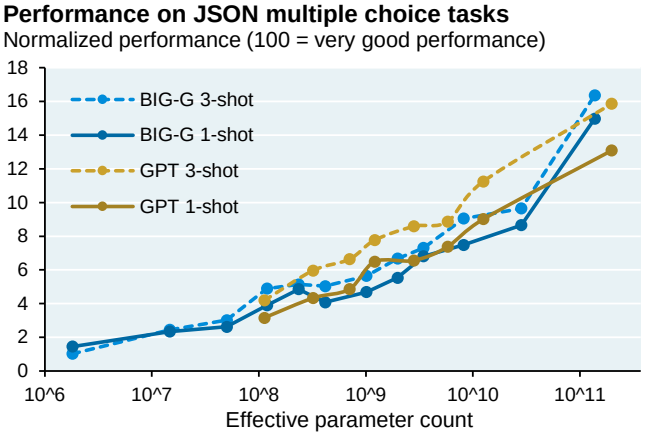
to solve, and I list a few of them in the Eye in the Markets.

Some of them are interesting. They ask the models to guess what movie they're referring to when they describe the plot just written in emojis. They have a take-the-GRE exam reading comprehension section where sometimes the right answer is more than one. They read short stories about crime and have to identify who the perpetrator was and the reasoning behind it. They have to see whether they can identify sarcasm or dark humor. You get the point. These are common-sense reasoning capabilities that are often beyond what the current models can do.

And they released the results last summer, and there's a long way to go. So the aggregate score that Google and OpenAI were getting at the time was about a 15 out of 100, where the average human was a 45, and the best humans were 70 to 80. So these models have a long way to go in understanding things and being able to put things in proper context. The larger the models get presumably, they'll be able to get better at doing this.



Source: Big Bench project. June 2022.



Source: Big Bench project. June 2022.

But these large language models, the battles are just beginning. A lot of times their capabilities are overstated. But it will be interesting to watch the narrow, more well-defined tasks that they get asked to do, 'cause that's really where the big productivity benefits are going to come from. And my feeling is they'll probably end up boosting productivity of companies rather than putting legions of people out of work. But we will see. Anyway, take a look at this month's Eye on the Market, and we get into all of these details and more. Good talking to you, and we'll see you next time, bye.

FEMALE VOICE: Michael Cembalest's Eye on the Market offers a unique perspective on the economy, current events, markets, and investment portfolios, and is a production of J.P. Morgan Asset and Wealth Management. Michael Cembalest is the Chairman of Market and Investment Strategy for J.P. Morgan Asset Management and is one of our most renowned and provocative speakers. For more information, please subscribe to the Eye on the Market by contacting your J.P. Morgan representative. If you'd like to hear more, please explore episodes on iTunes or on our website.

This podcast is intended for informational purposes only and is a communication on behalf of J.P. Morgan Institutional Investments Incorporated. Views may not be suitable for all investors and are not intended as personal investment advice or a solicitation or recommendation. Outlooks and past performance are never guarantees of future results. This is not investment research. Please read other important information, which can be found at www.JPMorgan.com/disclaimer/EOTM.

[END RECORDING]