EconoFact Chats: On the Economics, Geopolitics and Technological Challenges of AI
Chris Miller, The Fletcher School, Tufts University
Published on: February 16th, 2025

Michael Klein
I'm Michael Klein, executive editor of EconoFact, a non-partisan, web-based publication of the Fletcher School at Tufts University. At EconoFact, we bring key facts and incisive analysis to the national debate on economic and social policies, publishing work from leading economists across the country. You can learn more about us and see our work at www.econofact.org.

Michael Klein
In 1957 the Soviet Union launched the first satellite to orbit the earth. This "Sputnik moment" raised fears that the Soviets were going to win the technological war against America. The United States government began to fund research that led to great technological advances including, a dozen years after Sputnik, landing men on the moon. Now, a half-century later, the technological frontier is artificial intelligence. In January there was a new Sputnik moment when the Chinese company DeepSeek announced that it had the ability to conduct artificial intelligence at a fraction of the cost of competitors like ChatGPT. In the immediate wake of this announcement, the market value of the chipmaker NVIDIA lost a half-trillion dollars. What is the future of AI? What are the production, technological, and political challenges facing artificial intelligence development? And will AI live up to its promise of making life easier and better, or will it destroy jobs and potentially run amok? To answer these questions I'm very pleased to welcome back to EconoFact Chats my Fletcher colleague, Professor Christopher Miller. Chris is the author of *Chip War*, the book that won the 2022 Financial Times Book of the Year Award. Chris, thanks for joining me once again on EconoFact Chats.

Chris Miller
Thank you for having me back.

Michael Klein
Chris, what's happened with AI since you finished writing Chip War? It seems that there's been a series of amazing advances in just the last few years.

Chris Miller
Well, that's right, Michael. I think since the release of ChatGPT in November 2022, there's been a realization that progress in large language models has produced some really incredible new capabilities. And that, in turn, has driven a surge of investment in both building new AI models and developing new products based on them, and also in constructing huge quantities of the AI infrastructure that's needed to run these models, which means very large data centers full of the most advanced chips.

Michael Klein
So, talking about chips, of course, *Chip War* is about that, and you did a wonderful job in that of explaining the history of semiconductors and how they're made today. And it's been said that semiconductors may be, to the 21st century, what oil was to the 20th century. Do you think of this as a valid analogy?

Chris Miller
I think it is true, to the extent that today it's really hard to find almost anything that doesn't rely on computing power, which means semiconductors. It's true for dishwashers. It's true for medical devices. It's true for cars. And it's going to be increasingly true, as we begin to deploy artificial intelligence capabilities to a wider array of goods and services, all of which will require more and more advanced semiconductors. So, the pervasiveness of computing today is like the pervasiveness of energy as a factor of production in the industrial age. I think the second reason why there's some similarity between oil and chips is that the production is concentrated. In fact, it's even more concentrated in chips than it was in oil. Saudi Arabia produces 10% or 15% of the world's oil. Taiwan produces 99% of the GPU chips that make artificial intelligence possible. This concentration has both really significant economic, but also geopolitical implications.

Michael Klein
And of course, Taiwan is under threat because China considers it a breakaway province.

Chris Miller
Well, that's right. And that is one of the key risks hanging over both Taiwan itself, but also all of the world, because we collectively rely on TSMC, the Taiwan Semiconductor Manufacturing Company, to produce a very large share of the world's ships and an extraordinarily large share of the most advanced ships. And some of the biggest tech companies in the world, like NVIDIA, like Apple, they source almost all of their critical semiconductors from TSMC, which does the actual manufacturing of the chips that they and their products depend on.

Michael Klein
So, as I mentioned, Chris, in *Chip War*, you did this really wonderful job of explaining semiconductor fabrication. And you're a historian, so I was kind of surprised by that, but you're also really smart, so I guess that helps a lot. I'm going to ask you to do the same thing with artificial intelligence. Many of us have used ChatGPT or some other program, and you and I both hope our students are not writing their papers that way. But can you explain how semiconductors fit into the creation of machines that do artificial intelligence?

Chris Miller
There are two reasons why you need semiconductors and computing power in AI. The first is for training models. And today, when it comes to large language models, to simplify a bit, they're essentially trained by reading most of the text on the internet and identifying patterns in this vast quantity of words that they're reading. And identifying patterns helps them understand what connections they should make when asked questions in the future because they've identified the correlations between different uses of words in the past. And so, to crunch all of the data on the internet and to identify these patterns takes an extraordinary amount of computing power. And that's where companies like NVIDIA, which is the leader in the production of AI chips, has really enabled the training of these very powerful models. The second use of chips is in the deployment of these models, because it's increasingly the case that when you ask a question of a model, they take a fair amount of time to ‘think’ about the answer they're going to give you. And that thinking is the model working through what the optimal answer to your question is based on what it learned during the training process. And this also takes a lot of computation, which is why, for example, if you ask ChatGPT a hard question, and especially if you use one of the

newer reasoning models, it'll take 30 seconds or a minute to think about the answer before it spits it out. All of that is semiconductors working very, very hard for you while the AI model produces its answer.

Michael Klein
NVIDIA started out as a gaming company, didn't it?

Chris Miller
Yep, that's right. And NVIDIA realized in the 2000s that the same mathematics that was used to display images on computer screens or in video games was really pretty similar to the math that was being used in AI systems. And so NVIDIA started investing very early on in creating separate software that was designed to make it easy to write machine learning and AI systems and to run them on these gaming chips. And for a long time, AI was a secondary business segment for NVIDIA. But over the past couple of years, it's been the primary driver of the company's growth. And it's the main reason why today NVIDIA is the world's most valuable company.

Michael Klein
Well, as I mentioned in the introduction, DeepSeek seems to have bypassed the need for the chips like those created by NVIDIA. And as a consequence, NVIDIA's stock price fell by 17% over the course of a day or two. What does DeepSeek, a Chinese company, claim that it can do?

Chris Miller
Well, DeepSeek made it into media headlines and certainly attracted the attention of stock market investors in the middle of January. It released a new reasoning model similar to the o1 model series released by OpenAI. And DeepSeek claimed that it trained this model on a very small budget of around $5 million. And it also offered access to using its model for very low prices. And I think the media interpreted this as a surprising advance in artificial intelligence. In fact, I think for AI researchers, the surprise was that everyone else was taken by surprise by DeepSeek. In AI circles, DeepSeek was a matter of regular conversation throughout the latter half of 2024. I was really taken aback when my family members and friends were asking about DeepSeek all at the same time in January, given that none of the capabilities that emerged in January were a surprise relative to what they'd previously demonstrated. And I think that the key mistake that the media made was to conclude that DeepSeek was able to train a sophisticated model on a very limited budget. And I think what we know now is that that's not really true. They defined their training costs in the most narrow possible way and didn't include the rest of the research and development costs. And so the theory that it only took them $5 million to train a model is just not right. It probably took them one or two orders of magnitude more money, which would put them roughly on par with other big AI labs like Anthropic or OpenAI. So the headline, ‘DeepSeek trains model on a shoestring budget,’ I think is basically not true.

Michael Klein
I've read that if you ask DeepSeek a question about Tiananmen Square, it doesn't seem to answer those. Is that correct?

Chris Miller
That's true. DeepSeek, like a fair number of Chinese models, are censored for political reasons. I think we shouldn't, though, conclude that it's not a capable model. I think the censorship is probably fairly narrow. It won't talk about Xi Jinping. It won't talk about Tiananmen Square. It's cautious in talking about Winnie the Pooh. But that doesn't mean that for 99% of use cases, it won't give you a pretty good answer. I think it will give you a pretty good answer, and it does so relatively efficiently. The other aspect of DeepSeek that I think is worth looking at is the cost it takes to use it. And here, there's a really interesting question, which is, is DeepSeek charging at cost for use of this model? Is it charging with a bit of a margin? Or is it even offering use below cost in an effort to win market share? The strongest version of the argument is that DeepSeek might be doing to AI what China's already done to other industries like solar panels: provide subsidized production, lower cost in an effort to win market share, and imperil the business models of foreign firms. I think that's an interesting issue to track. It might be related to the fact that when DeepSeek was first launched, it actually had such demand for its services, it was unable to service them all and had to shut off access to non-Chinese users, which suggested, again, that they weren't actually pricing at a level that they were able to sustain in the long run.

Michael Klein
I want to go back to that idea of censorship. You said only a small set of things are censored. But if I think about what I talk about during the day with my wife or my friends or my students, only a small part of that is politically sensitive. So there's no reason to censor when I'm asking what we're having for dinner. But if I start to talk about, say, the geopolitical situation, that could get censored. So it seems like there's some real challenges with this, almost, you know, sort of Orwellian challenges with this going on. Would you agree with that?

Chris Miller
Yeah, I think that is right. I think we're entering, in general, a new era in which we no longer search for information on the internet. We ask the internet and these AI models to give us information. And this is a particular challenge in authoritarian countries with censorship, like China, where you're going to have the government actively saying to companies, do this, don't do that. But I think it does pose general questions in addition, because unlike Google, where if you search for something, it'll give you 10 different search results in chat GPT, ask a question, it'll give you one result. Or if you ask for two results, it'll give you two, but only if you ask for it. And so I think this will pose an interesting set of dilemmas and trade-offs for us as a society, as we increasingly learn to trust AI models as arbiters of truth. And I think we'll have to collectively develop an ability to parse what models are telling us and learn to assess when they're likely to give us accurate versus inaccurate answers.

Michael Klein
So maybe not just, you know, explicitly autocratic countries. If you think about what happened in the last few months in the United States, for example, Jeff Bezos prevented the endorsement of Kamala Harris from the Washington Post, which he owns. And there seemed to be sort of a bending of the knee of tech giants to the new administration. So is it conceivable to you that this kind of thing could happen with American platformed artificial intelligence as well?

Chris Miller
Well, I would say there's inevitably going to be a relationship between the politics of the day and the outputs of AI models. And that's because humans play a role in training AI models. They play a role in helping models prioritize what to focus on. They play a role in the quality and safety evaluations that are done at the end of the training process. And so to the extent that you prioritize or deprioritize anything that will be present in the model. And so I think you see this from both sides of the political spectrum. The right alleges that models are being too politically correct, for example. And I think you're seeing now with political change in the last couple of months concerns on the left that you'll have models being biased in a more right-wing direction. I think there's no single way to be sure models aren't biased. They're inevitably biased by the information you train them on, which is in turn a feature of the choices that humans make when they set up the training. And so I think that speaks to the benefits of having competition, multiple models to choose from. But it also speaks to the importance of teaching individuals to be discerning users, just like we want people to be discerning readers of the news, discerning browsers of the internet. We'll also need to become discerning users of models. So we learn what we do trust and what we do not.

Michael Klein
Chris, you already mentioned TSMC, the Taiwan Semiconductor Manufacturing Company, and how it makes 99% of the chips that are used for AI. So I suppose this adds another dimension to the U.S.-China conflict, since, as I mentioned, the Chinese Communist Party regards Taiwan as a renegade state. And it's been very clear about its views that Taiwan should be part of the People's Republic.

Chris Miller
Well, that's right. And that is a key challenge to the future of the AI industry. It's already the case that over the past couple of years, leading tech companies, leading AI labs have struggled to get access to enough of the GPU chips and the high bandwidth memory needed to build AI servers and train and deploy AI systems. But it's particularly challenging in the future, potentially, because of the concentration of production of these chips in Taiwan and to a lesser extent in Korea as well. And so there's a really direct relationship between peace in East Asia and progress in AI, because the way supply chains are structured today, U.S. tech companies, world tech companies couldn't get the chips they needed without access to Taiwan. And I think it's worth lingering on the extent to which Taiwan is just at the absolute epicenter of AI supply chains. We often think of AI models as being disembodied, something that exist in the internet. But in fact, the physical products that make AI possible depend on supply chains that, to a very large extent, get traced back to a single country in East Asia.

Michael Klein
So would things have been a lot easier if the advanced chip technology had been developed in Zurich?

Chris Miller
Well, I think had that been the way the chip industry developed, we'd be a lot less nervous today. But I think it's also the case that Taiwan has now developed some pretty strong comparative advantages in chip making in terms of its workforce, in terms of the experience that's built up

there, in terms of the entire supply chain players, from the chemicals to the materials to the technicians that can repair the tools when they're broken. All of that expertise is concentrated in and around Taiwan. And so it would be very difficult now to shift the industry to Zurich, even though it might be nice had that been the way it developed.

Michael Klein
Well, how about shifting the industry to the United States? One potential way to address the challenges is to onshore production of semiconductors. The 2022 CHIPS and Science Act, where the acronym CHIPS stands for Creating Helpful Incentives to Produce Semiconductors, was meant to bring microchip manufacturing to the United States after decades when the fabrication of semiconductors moved to, as you mentioned, places like Taiwan and South Korea. How successful has the CHIPS Act been so far, and what are its prospects for success in the future?

Chris Miller
So I think the goals of the CHIPS Act were, first off, to boost investment in chip-making facilities in the United States, and secondly, to encourage cutting-edge R&D into chipmaking technologies to guarantee the U.S. technological lead when it comes to as many segments of the chip supply chain as possible. If you look at the first goal, boosting investment, we've seen a dramatic increase in investment in chip-making facilities in the U.S. for the past couple of years since the passage of the CHIPS and Science Act relative to the prior trend. Multiple times more spending in recent years in terms of capital investment in new facilities than you saw in the prior decade. And that's, I think, directly due to the subsidies as well as the tax credits that Congress offered via the CHIPS Act. So measured purely in dollar terms there's been a huge increase. Congress has gotten what it wanted. I think, though, there's some challenges in terms of ensuring that these dollars invested are resulting in the types of output that you'd optimally want to guarantee your economic security. One of the challenges is cost. It's 30 or so percent more expensive to make chips in the U.S. versus Taiwan. And so even if you're investing more, you're getting fewer chips per dollar of investment than you would if you'd invested a dollar in Taiwan. But second, there's a lot of different types of chips. And when it comes to the most cutting-edge chips, the most advanced, they're still only made in Taiwan. And so companies like Apple and NVIDIA, they're sourcing some chips from the U.S., but they're sourcing the most important chips still from Taiwan. Second question is on cutting-edge R&D. And here, I think, we're still in the early stages of the CHIPS Act beginning to set up the R&D facilities, the pipelines of talent from academia to industry. But I'm relatively optimistic that these institutions will have a major impact in solving the key challenge in the chip industry, which is moving ideas from the lab and universities towards the fab where they're actually brought into production.

Michael Klein
So, Chris, you're trained as a historian. You've written this book about advanced technology. But I know you also have a really strong background in economics. So I'd like to ask your views on what a world with pervasive AI would look like. Any type of technical change is associated with both the destruction of certain kinds of jobs and the creation of other new types of jobs. There aren't many TV repairmen anymore, but there are people who create games and so on. So overall employment might not be affected, but some groups of people will benefit and others will lose out. And acknowledging how hard it is to forecast these kinds of things, I'm still going to put you

on the spot and ask for your opinion about what the economic landscape will look like as AI becomes more pervasive.

Chris Miller
Well, I think one of the key areas of uncertainty in answering this question is the rate of change. We've always had technology improving, different ways to measure it. But I think if AI improves at roughly the rate that computing technologies have improved in general over the past half century, we shouldn't expect the social implications to be more disruptive, nor the labor market implications to be any more disruptive. And so given that that would be my rough base case, I think we're probably overestimating the scale of labor market turnover or turmoil that we're going to see. And I think there are lots of examples from economic history of expectations of large-scale unemployment that really weren't matched in reality. My favorite example is of employment in bank branches. When the ATM was first invented, there were widespread predictions that banks would need fewer employees because they wouldn't be handing out cash. And today, there are more people working in banks because they're selling mortgages and other types of products. And so I think that basic trend line is probably likely to hold. I think the concern I would have is if we have much more rapid progress than prior trends, in which case it'll be harder to digest the labor market impact. But that could be a good scenario as well as a bad scenario. If, in fact, we reach an age of rapid productivity growth, that will create much more wealth. And then the question is just, can it be allocated in a way that addresses the labor market impact that's created? But right now, I don't think we see evidence that we're having this extraordinary acceleration in productivity growth. It's certainly not present in economic data. And so for now, I think the most reasonable assumption to make is that technology will continue to improve. We've reached the new phase of computing technology, which is AI. And it's going to be one of the key drivers of productivity improvements, but not a revolutionary one, or at least not any more revolutionary than smartphones or PCs or mainframe computers were in their day.

Michael Klein
Or electricity?

Chris Miller
Or electricity. I mean, one of the fascinating dynamics is if you go back and look and try to, in the economic data, find the productivity impact of electricity or computers, it's actually harder to find than you might think. Anecdotally, it's easier to see. In the data, it's harder to find.

Michael Klein
All right, Chris. So that's your economic hat. I'm going to ask you to put on your historian's hat for a moment and address the question of the deeper concern that in a world with pervasive AI, things could just run amok. Some people talk about real apocalyptic possibilities. What do you think about that?

Chris Miller
Well, I think it's impossible to fully reject or disprove apocalyptic scenarios. But I think, I guess I'm struck by our ability to develop institutions to manage risks in other technologies over time. There's lots of other apocalyptic technologies that we've created. We've just learned to use them in ways that aren't that destructive. And the fact that we've now for 70 years managed to live in a

world without nuclear weapons use, even though we have nuclear weapons, I think is a positive sign. I'm much more worried that we slow down AI progress by wrongly fixating on the worst case scenario. And I think that there's so much concern in the AI community about safety that it actually misses, that although safety is important, quality is even more important. And we should be focusing just as much on ways to make sure we're using AI in ways that are going to improve outcomes in healthcare and in the provision of consumer services and in government services, for example. So I worry much more about our slow adoption than I worry about AI displacing or replacing or causing humans to go extinct.

Michael Klein
Well, that's a very optimistic view, and it's nice to be able to end on an optimistic note rather than the kind of pessimistic notes that one often hears in these discussions of artificial intelligence. So, Chris, thanks for joining me once again on EconoFact Chats and shedding light on a very complicated technical subject and doing so with a lot of clarity.

Chris Miller
Thank you again for having me.

Michael Klein
This has been EconoFact Chats. To learn more about EconoFact and to see the work on our site, you can log into www.econofact.org. EconoFact is a publication of the Fletcher School at Tufts University. Thanks for listening.