

April 2025  
Geoff Huston

## Jevons Paradox and Internet Centrality

The story of computing and communications over the past eighty years has been a story of quite astounding improvements in the capability, cost and efficiency of computers and communications. If the same efficiency improvements had been made in the automobile industry cars would cost a couple of dollars, would cost fractions of a cent to use for trips, and be capable of travelling at speeds probably approaching the speed of light!

We've been the beneficiaries the bounty generated by Moore's Law for most of this period in the evolution of chip fabrication techniques, where, since the late 1950's the transistor density of integrated circuits has been doubling every 18 months or so (Figures 1, 2). At the same time the fabrication cost per transistor has fallen over much of this period (Figure 3). Its only in recent years as the chip industry moves to smaller than 7nm feature technology has the per gate cost stabilised. However, during this same period the number of transistors per chip is rapidly approaching 1 trillion transistors.

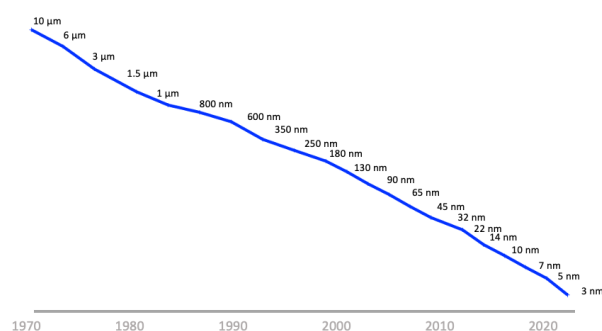


Figure 1 – Silicon Feature size over time (From Wikipedia: 3\_nm\_process)

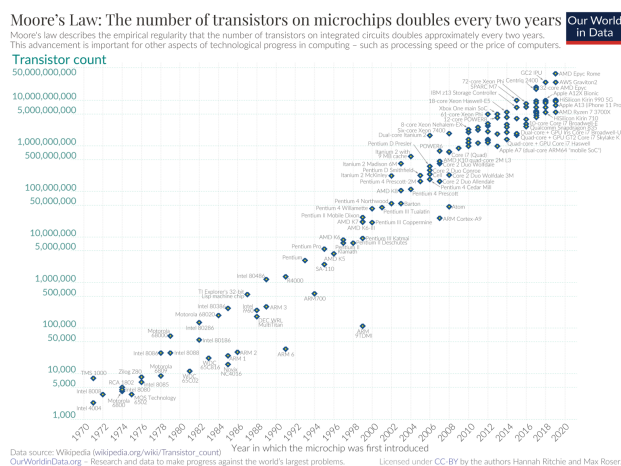


Figure 2 – Transistor count in microprocessors up to 2019 - By Max Roser, Hannah Ritchie - <https://ourworldindata.org/uploads/2020/11/Transistor-Count-over-time.png>, CC BY 4.0.

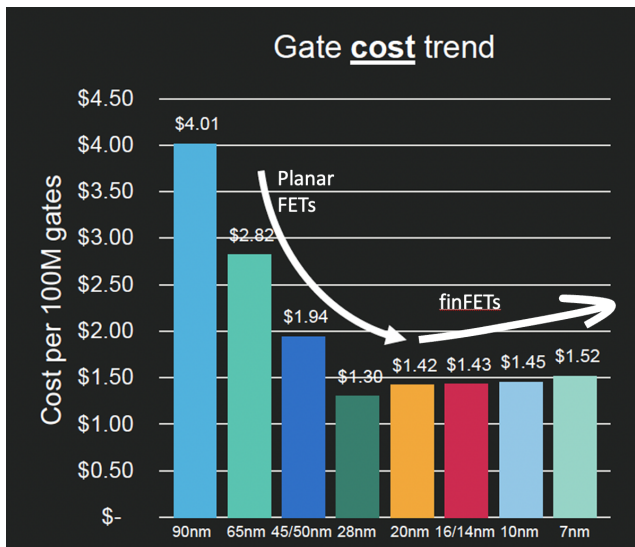


Figure 3 – Gate cost trend (From Marvell Investor Day 2020 presentation)

It has been more than gate density. For the past two decades or so, demands of the mobile device market forced chip designers to produce processor chips where the total energy demands of the device were able to be provided by a single battery charge for a full day's use, with the energy efficiency of the chip held the chip's heat to such levels such that it did not become literally too hot to handle! The aim was to increase the processing capability of the chip, while at the same time keeping the total energy consumption of the chip roughly constant!

How does all this relate to Jevons' Paradox?

### William Stanley Jevons



William Stanley Jevons was one of the founders of neoclassical economics in the mid-nineteenth century. He witnessed the period of railway mania the 1840's which saw a huge speculative bubble fund the construction of track, engines, carriages, and the expansion of coal mining.

Steam locomotion made the transportation of goods and people cheaper and faster. One of the impacts of this was the rapid expansion of coal production in Britain, where pre-railway production volumes of some 20M tonnes of coal per year rose to 120M tonnes per year by 1860 and was on a trajectory to further increase in the coming years. In 1865 Jevons published *The Coal Question* that speculated on the likelihood of exhaustion of British coal deposits as a result of the rise of coal-fired steam engines. One of the key observations in his book was that, paradoxically, the consumption of coal actually increased when technological progress improved the efficiency of steam engines. We tend to believe in the concept that improvements in efficiency in the use of a resource would naturally lead to a drop in the level of resource consumption, as the same level of activity would consume less of the resource. In the situation where there is a limit to the volume of production of the resource, then these improvements in the efficiency in the use of a resource should get us to a point where total consumption of the resource lies within "natural" production volumes, such the resource is being used at a *sustainable* level.

Jevons Paradox offers a diametrically opposite view, where improvements in efficiency of resource utilisation acts as a positive incentive to increased resource consumption, and the net result is a increase in resource consumption. Efficiency lowers costs, which lowers prices, which increases

demand. And, sometimes, this increase in demand is disproportionately large, with the result that overall consumption actually grows. This outcome came to be known as the Jevons effect, or Jevons paradox. The pursuit of increased efficiency does not inevitably lead to sustainability.

He observed that ways to increase the efficiency of coal-fired steam engines would not necessarily suppress the demand for coal, but, paradoxically, may well stimulate even greater levels of demand for coal.

As it turned out, Britain did not in fact run out of coal reserves. Coal production peaked in 1913 at some 280M tonnes per year, and then declined over the ensuing decades to total production levels not seen since the 1700's.

The answer to this seeming paradox lies in a concept termed the *price elasticity of demand*. When more efficient coal engines lowered the cost of coal-powered transportation of people and goods, the total demand for transportation increased by far more than the gains to be made from increased efficiency. In more general terms in a *price elastic* scenario, any increase in the resource utilisation efficiency of a resource will lead to a drop in the unit price of the good or service that is derived from the resource, and the lower unit price of the good or service may lead to increased demand.

## Computers and Communications

There is a similar story in the evolutionary path of computers and communications systems.

One of the earliest electronic computers, ENIAC, was constructed in 1945. The 30 tons of equipment housed some 18,000 vacuum tubes and consumed 150kw of power. The inventions of the transistor in 1947, and the integrated circuit in 1959, did not result in smaller more efficient version of ENIAC, but instead these advances facilitated the development of vastly more capable processing systems that operated at significantly greater level of efficiency, yet in absolute terms consumed far more energy than the 150kw of ENIAC. For comparison, Nvidia's A100 chip has a constant power consumption of 400w, and the latest H100 consumes some 700W. A large AI-capable datacentre with one million of these GPUs would be looking at a power consumption of around 1.5GW. Yes, they are far more energy efficient per single operation than ENIAC, but this increased efficiency has led to vastly increased number of operations per second and a far greater level of energy consumption.

A similar picture can be seen in the evolution of Apple's iPhone. The first iPhone, released in 2007, featured a Samsung 32-bit ARM microprocessor with a 412Mhz clock, 128MB memory and a 3.7V 1,400mAh lithium-ion battery. Some twenty years later the iPhone 16 uses a pair of 64-bit ARM processors (using 3nm fabrication processes) with a 4Ghz clock, 8Gb of memory and a 3.89V 4,685mAh lithium-ion battery. That's approximately three times the battery capacity for well over 100 times the performance.

These ARM processors have been used across a range of Apple devices, and the relative performance of these processors over is shown in Figure 5. All these improvements in the hardware of the device, including processing power, battery technology and energy efficiency has not resulted in a cheaper device with roughly constant performance parameters, but in significantly greater levels of capability that is intended to render superseded models obsolete and stimulate demand for the current product.

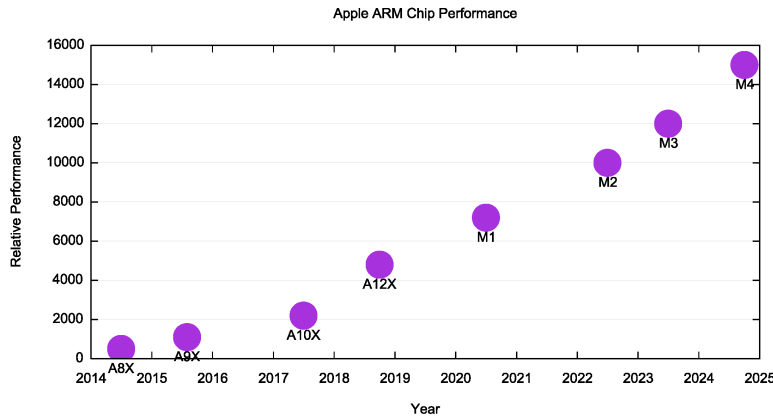


Figure 5 – Time Series of Apple ARM chips

On a more general level, the efforts of the silicon chip designers and foundries to produce better chips, with progressively increased performance and greater energy efficiencies has not reduced the requirements of devices in terms of their size and energy requirements but conversely has increased device capabilities and stimulated higher levels of overall demand. The same appears to be the case at the macro level when we look at the requirements of data centres, in terms of their processing capability, their energy requirements, cooling capacity and footprint.

This is also evident in the application space. An example lies in the adoption of blockchain-based digital artifacts, such as Bitcoin and its derivatives, Non-Fungible Tokens (NFTs) and blockchain-based name systems. The increase in power and efficiency of computing platforms enables the viability of these services, but the ensuing service uptake may well increase total levels of resource consumption well beyond any potential gains that could be made through increased efficiency.

It has been estimated that Bitcoin mining in Iceland uses around 120 MW of power, or around 85% of the 140 MW of power used by the country's data centres in 2022. Iceland's abundance of renewable energy and cheap power has had both data centres and Bitcoin mining operations flocking to the country in recent years to set up shop. Iceland's cool climate is another benefit, as data centres produce a lot of heat that would require additional energy to cool if located in a warmer climate. Both politicians and environmental activists have questioned the benefit of Bitcoin and digital currency mining operations for the Icelandic nation as well as their impact on the environment. The article that reported this estimate also carried the comment that "This is a waste of energy that should not be happening in a society like the one we live in today."

<https://www.icelandreview.com/news/nature-travel/bitcoin-mining-a-growing-waste-of-energy-in-iceland>

### The Jevons Paradox and Centrality

The Jevons paradox is counterintuitive because, other things being equal, we would normally expect higher efficiency to cause lower resource use. The Jevons paradox occurs, however, precisely because other things are not equal. What is different is demand. Although Jevons himself

could not have known about the concept of price elasticity of demand, an economic concept whose study which still lay decades in the future at the time he raised the issue, he did anticipate the essence of that idea. He also recognized that as energy became cheaper, the total rate of economic growth across society would increase as well, because energy is an input to virtually all other goods and services.

In a narrow technical sense, the Jevons paradox only happens if a technology makes an existing process more efficient, and only if demand is highly price sensitive. But new technology often obviates the old way of doing things entirely. The Jevons paradox might be taken to mean that new technologies always create new problems of their own. For example, combustion engine vehicles might have solved the environmental problem of copious quantities of horse manure in metropolitan environment, but they created their own problems of air pollution and greenhouse gas emissions. (Never mind the fact that we would need more than 10 trillion horses to move freight and passengers as many miles as we now do with vehicles!).

But can we explain today's headlong rush into AI using Jevons Paradox alone? Is it the outcome technological evolution crossing a threshold of price and capability that has unleashed formerly pent-up user demand? Did we always want AI-based tools in our IT environment but could not afford to access such tools until now? Or are other pressures at play?

The disturbing characteristics of today's digital environment is the dominant position of a small clique of digital behemoths whose collective agenda appears to be based more on the task of ruthless exploitation of everyone else in the singular pursuit of the accumulation of unprecedented quantities of capital and social power. The continual pressures of technical evolution through the application of Moore's Law to chip fabrication have a lot to do with this. The underlying platform of every digital service is one that will likely be twice as powerful, and, paradoxically, half the price of the current platform in as little as two years. This would conventionally force providers to constantly upgrade their platforms to stay ahead of their potential competition. There is another path, however, that is based on complete market domination. Such a position allows the dominant incumbent to shape the market according to their own desires, and push product and services to consumers at price points that are inaccessible to any competitors. This "winner take all" form of supply-side push in markets is one that reinforces the position of dominant incumbents.

Are the large digital service entities engaged in a huge investment program to build AI datacentres simply responding to an unprecedented demand due to huge volumes of demand for AI-generated content and service? Or are they motivated by a desire to occupy this space before any potential competitor can amass sufficient volumes of capital and momentum to effectively challenge the position of the incumbents?

I strongly suspect we are in the latter situation. Jevons Paradox appears to apply in demand-pull markets which are price elastic. Marginal efficiencies gained in the production of the good or service are reflected in unit price drops, which unleashes a disproportionate level of further demand. What we appear to have in today's highly centralised digital space is supply-push dynamics, where the current incumbents push novel goods and services onto the market at a price point that bars any new market entrant from competing. These supply-side push markets work to not only pre-empt future demand by anticipating it far in advance over any potential competitor, but to influence the nature of demand that that consumers are inured to desire the service being offered. This circularity works to further reinforce the position of the incumbent providers.

It's challenging to see a way through the current situation. Regulatory efforts to disrupt the operation of these large providers are often either ineffectual or risk voter displeasure, while a

laissez faire stance on the part of public regulatory agencies achieves little other than ensuring that the current incumbents maintain their entrenched position for many decades to come.

But this is not a stable situation. Physics has a lot to say when the feature size of elements in a silicon wafer get down to around 1nm in size. Further reductions in the size of what is essentially a planar process can't be achieved due to physical behaviours such as electron tunnelling. Our current efforts to address this has been to introduce three dimensional structures onto the chip, such as FinFETS, and to explore other materials instead of silicon as the chip substrate, including Gallium Nitrate, and the somewhat hazy prospect of quantum computing. But the success of these efforts in achieving further radical improvements in the size, complexity and energy efficiency of computer chips is by no means assured. If we cannot keep on doubling down every two years or so in the production of processing chips then the technological innovation pressures on the industry will inevitably weaken.

It's a bit like the childhood game of musical chairs, where the aim is to claim your seat when the music stops. The music of technical innovation in chip fabrication is still playing, but it's clear that its tempo is slowing down. What happens after the music stops remains anyone's guess, but the conventional thinking is that at such a time you really should've claimed your seat, and this observation is probably uppermost in the current thinking of the digital behemoths.

---

## Disclaimer

The above views do not necessarily represent the views or positions of the Asia Pacific Network Information Centre.

---

## Author

*Geoff Huston* AM, M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region.

*[www.potaroo.net](http://www.potaroo.net)*