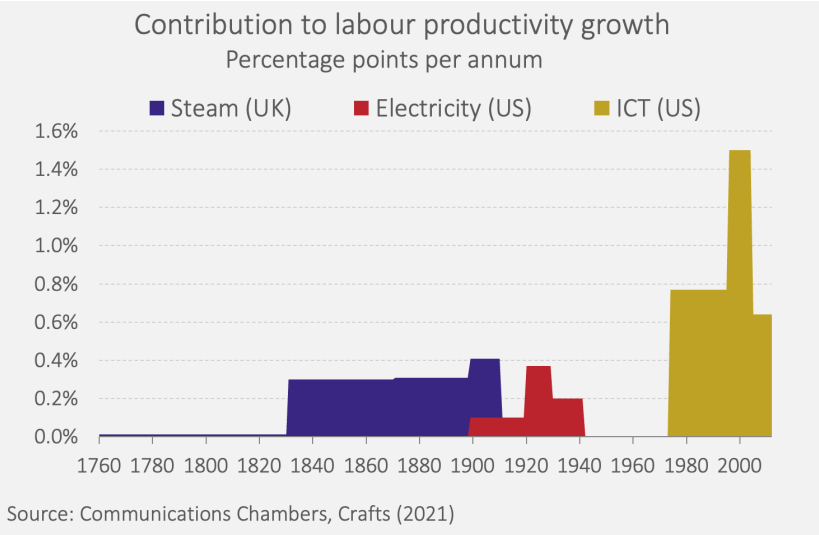
AI governance – lessons from other general-purpose technologies

Progress

"Productivity isn't everything, but in the long run it is almost everything." Paul Krugman, 1992

Technical progress underpins productivity growth, and productivity growth is the only sustainable source of growth in income per capita. Over the past two centuries productivity growth has increased income, leisure and contributed to longer life expectancy. Productivity growth, whilst involving a reallocation of labour, has not harmed overall employment.

General purpose technologies such as steam, electricity and connected computing underwent continuous improvement, saw widespread adoption, and made substantial contributions to productivity growth.[1]



The figure above illustrates how the productivity growth contribution of steam took a long time to materialise since steam engines were initially inefficient and because it took time for complementary innovations such as factories and railways to be developed. Electricity saw more rapid adoption, whilst computers had little impact until networked computing was adopted.

Artificial Intelligence (AI) is anticipated to be a general purpose technology which drives a new wave of productivity growth. AI is also the invention of a method of invention, which is expected to contribute to breakthroughs in areas including health. A renewal of economic and social progress due to the application of AI would be welcome, and may also help mitigate zero-sum thinking, a symptom of low growth.[2]

Previous general purpose technologies were not regulated via specific legislation - we did not have a law of steam, electricity or connected computing. Nor were applications subject to assessment requirements prior to release. Innovation without permission was an enabler of progress.

Applications of previous general purpose technologies were of course subject to existing regulation and new regulation was developed where issues arose e.g. rail safety in relation to an application of steam.

[1] Nicholas Crafts, September 2011, Artificial intelligence as a general-purpose technology: an historical perspective, Oxford Review of Economic Policy, Volume 37, Number 3. https://academic.oup.com/oxrep/article/37/3/521/6374675
[2] Sahil Chinoy *et al*, September 2023, Zero-Sum Thinking and the Roots of U.S. Political Divides. https://www.nber.org/papers/w31688#fromrss

We should not abandon the principle of innovation without permission in relation to AI, rather we should rely on adaptation of markets, existing law and regulation, and only legislate to address targeted issues where these adaptations prove insufficient.

Uncertainty

> *"The coming of the wireless era will make war impossible, because it will make war ridiculous."* Guglielmo Marconi, 1912

Technical progress is itself uncertain (though regularities such as Moore's law may hold for a time). Further, uncertainty regarding the economic and social impacts of technical progress are compounded by two considerations.

First, progress tends to involve complementary innovations, which almost by definition are non-obvious, for example electrification allowed the redesign of factories with distributed power and the adoption of new production processes.

Second, the consumer response to new goods and services made possible by new technology is uncertain and consumers themselves may contribute to the outcome, for example, via user generated content.

The above two sources of uncertainty do not flow directly from the path of the technology itself, so those who understand the technology may have no advantage when predicting the implications. Indeed, their technology focus may blind them to the implications.

What happens in practise is that entrepreneurs test alternatives via the market, and some succeed. Those providing capital also prioritise investment opportunities they anticipate are most likely to succeed. Both mechanisms are decentralised and competitive means of discovering beneficial applications. They are key mechanisms for the alignment of AI with our preferences.

Novelty

> *"Newsreaders still feel it is worth a special and rather worrying mention if, for instance, a crime was planned by people 'over the Internet.' They don't bother to mention when criminals use the telephone or the M4, or discuss their dastardly plans 'over a cup of tea,' though each of these was new and controversial in their day."* Douglas Adams, 1999

A recurring theme with new technology is that existing (and new) harms tend to be attributed to the technology *per se*. Competitors may also invent or exaggerate harms to gain advantage, for example, as occurred during the "war of the currents" in the late 1880s and early 1890s over whether to rely on alternating or direct current for electricity.

Harms also tend to be considered in absolute terms rather than relative to the harm arising with existing technology (for example, focusing on accidents involving autonomous vehicles, rather than the accident rate versus vehicles driven by people). We are in this phase with AI.

Capable technologies can be used for good or bad purposes. Further, it is likely that the greater the capability for good the greater the capability for harm. For example, a sharp knife is both more useful and more dangerous than a blunt knife. A safe knife would be a useless knife.

Large language models, and more generally foundation models with wide application, are a current focus in terms of novelty. An imperfect analogy might be a public library which is an open-source cultural repository with a wide range of potential applications. When the UK Public Library Act was debated in 1850 there was some opposition on grounds that libraries could give rise to unhealthy social agitation.

Further, under the European Union Artificial Intelligence Act the training data used in 'high risk' applications should be complete, unbiased, and free of any false information. Yet, considering the parallel with public libraries, these contain many books which include false information and bias, judged against our current state of knowledge and norms. We should pause for thought before contemplating removing almost the entirety of human knowledge and cultural history from the training data set for AI.

Finally, besides novelty, scale itself may be sufficient to qualitatively change the nature of harm. For example, connected computing brought a heightened focus in relation to data protection. What it did not bring forward, however, were calls for a law of connected computing. Likewise, our focus in relation to AI should be on specific harms, rather than attempting to write a "future proof" law of AI.

Neutrality

> *"The Americans have need of the telephone, but we do not. We have plenty of messenger boys."* Sir William Preece, Chief Engineer, British Post Office, 1878

A neutral stance as to how things are done is essential to allow progress. Regulation should therefore, to the extent feasible, be technology neutral. Neutrality is important for two reasons.

First, to ensure that the most efficient technology – human or machine or a combination – is utilised. If new regulation is applied to AI systems only, then the potential for productivity gains would be limited to the extent that the adoption of more efficient AI systems was discouraged. Second, to ensure that the safer option is chosen. If new regulation was applied to AI systems only, then the potential to reduce risk would be limited to the extent that the adoption of less risky AI systems was discouraged.[3]

This implies that AI should not be subject to additional requirements that do not apply to competing human decisions and systems. By default, absent new AI regulation, this will apply since services provided with AI inputs will be subject to existing regulation.

If new requirements are introduced, for example, related to the scale of harms associated with a particular AI application, then the new requirements should be applied to competing human systems. For example, to the extent AI raises concerns about surveillance, then any revision to limitations on surveillance should apply to machine or human based surveillance systems.

Adaptation

> *"Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"* Lewis Carroll, Through the Looking-Glass

[3] Categorising AI systems by risk, as proposed the European Union Artificial Intelligence Act, may exacerbate this problem. High-risk applications are typically already regulated and may correspond to the areas where the greatest gains from AI in terms of safety could be achieved, for example, in relation to transport and healthcare. Placing additional requirements on AI in "high-risk" categories may therefore delay safety improvements and reduce safety overall.

Before legislating to address any specific concerns in relation to AI applications we should consider the potential for adaptation.

Advances in AI are likely to be one of the key adaptive defences against adversarial systems. Holding back progress would therefore involve risks in relation to, for example, cyber security. An early form of adaptation was the use of AI to filter spam from e-mail.

More generally, markets can be expected to adapt to reduce harm and meet consumer's needs. This process of selection can include innovation in relation to market institutions which better assure consumers that something is genuine or safe.

Existing regulation can also adapt without the need for new legislation. For example, the process by which we ensure that drivers are sufficiently safe to drive on public roads needs to be adapted to incorporate autonomous vehicles. The ways in which safety is verified may differ, but the underlying aim should be the same.

Existing law may also need to be reinterpreted given new circumstances, or simply the scale of impacts with AI. An example is litigation in relation to AI and copyright. New legislation in relation to copyright may be required, however, it is also possible that the interpretation of existing law will prove sufficient.

Barriers

> *"Economic growth may be constrained not by what we do well but rather by what is essential and yet hard to improve."* Aghion, Jones and Jones, 2017[4]

A focus on removing barriers to AI should extend beyond the technology itself since complimentary innovation is a necessary part of the overall process. Existing interest groups, particularly those with licencing powers including professional bodies and regulators, may also seek to maintain the *status quo* by impeding innovation and progress.

The potential of AI may also be constrained by inputs, including energy and human capital. Energy is required to train and run AI models and to do the things in the physical world that additional intelligence may enable. We should seek to enable rather than constrain an expansion in energy supply if required - consistent with mitigation of climate change.

Human capital constraints tend to be self-correcting, and allowing skilled immigration can help. However, constraints on commercial and residential construction may constrain agglomeration effects, thereby limiting the potential of human capital, and contribute to inequitable outcomes in relation to the proceeds from productivity growth.

Finally, if we lose sight of those areas of the economy such as health, education and construction that have witnessed slow productivity growth we are likely to be disappointed by the aggregate economic impact of AI.

[4] Aghion, Jones and Jones, October 2017, Artificial Intelligence and Economic Growth. https://www.nber.org/papers/w23928

Ethics

> *"It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest."* Adam Smith, 1776

Ethical principles will apply to the application of AI irrespective of whether we explicitly "bring ethics to AI". Those developing AI applications want them to succeed in the marketplace and consumers will adopt those applications they find most beneficial. The market, involving free exchange, embodies the moral principle of mutual betterment consistent with self-interest. We understand that the market is incomplete due, for example, to externalities; so we complement it with targeted interventions which are underpinned by the ethical principle of Pareto optimality. Our political and legal institutions also address distributional concerns.

These interventions and considerations apply whether goods and services are produced utilising AI inputs or by other means. The development and success of AI applications will therefore be governed by ethical principles. That leaves open the question of whether fresh consideration of ethics would contribute to better outcomes. However, in considering this question, a comparative approach should be adopted considering the ethical framework governing existing markets and regulatory institutions.

If additional 'ethical' requirements, beyond those embodied in market selection and economic welfare-based interventions, delay AI that may also defer gains in income, safety and health (thereby resulting in avoidable deaths due to foregone medical advances) – that itself would arguably be unethical.

Policy

Maintain innovation without permission. Innovation without permission applied to previous general purpose technologies including steam, electricity, and connected computing; and was key to their success.

Recognise that advances in AI will be a key defence against adversarial AI. Holding back progress would involve risks in relation to cyber security etc.

Focus on adaptation rather than a 'law of AI'. We should rely on adaptation within markets, and in relation to existing law and regulation, and only legislate in relation to specific issues where these mechanisms are insufficient. Where existing governance is a barrier to AI driven transformation, we should reform it consistent with legitimate public policy objectives.

Seek regulatory neutrality between AI and human systems. Where AI applications are introduced into markets with existing regulation, they should be no worse than existing applications from a public interest perspective. Where regulatory changes are introduced, they should apply equally to AI and human systems. We should not discriminate against machines, otherwise we will forgo safety and productivity gains.

Brian Williamson[5], 2023

[5] Partner, Communications Chambers (brian@commcham.com). This paper represents the view of the author only, and does not represent a corporate view of Communications Chambers.