# Economic growth under transformative AI

Philip Trammell (Global Priorities Institute, Oxford University) and Anton Korinek (University of Virginia, NBER and CEPR)

Global Priorities Institute | October 2020

(last updated October 2023)

*GPI Working Paper No . 8-2020*

Please cite this working paper as: Trammell. P. and Korinek, A. (2020) Economic growth under transformative AI. *Global Priorities Institute Working Paper Series, No. 8-2020* Available at: https://globalprioritiesinstitute.org/philip-trammell-and-anton-korinek-economic-growth-under-transformative-ai/

GLOBAL PRIORITIES INSTITUTE

UNIVERSITY OF OXFORD

# Economic Growth under Transformative AI*

Philip Trammell[†] and Anton Korinek[‡]

October 2023

**Abstract**

Industrialized countries have long seen relatively stable growth in output per capita and a stable labor share. AI may be *transformative*, in the sense that it may break one or both of these stylized facts. This review outlines the ways this may happen by placing several strands of the literature on AI and growth within a common framework. We first evaluate models in which AI increases output production, for example via increases in capital's substitutability for labor or task automation, capturing the notion that AI will let capital "self-replicate". This typically speeds growth and lowers the labor share. We then consider models in which AI increases knowledge production, capturing the notion that AI will let capital "self-improve", speeding growth further. Taken as a whole, the literature suggests that sufficiently advanced AI is likely to deliver both effects.

*Thanks to Tim Blomfield and Nicholas Emery for contributions to the literature summaries and organization and Don Suh for research assistance.

[†]Global Priorities Institute and Department of Economics, University of Oxford. Part of this work was performed while the author was a Visiting Graduate Fellow at the Digital Economy Lab, Stanford University. Contact: philip.trammell@economics.ox.ac.uk.

[‡]Department of Economics and Darden School of Business, University of Virginia; Centre for the Governance of AI (GovAI); NBER; and CEPR. Part of this work was performed while the author was a Visiting Fellow at the Brookings Institution. The views expressed in this work do not necessarily reflect the views of the Brookings Institution. Contact: akorinek@virginia.edu.

# Contents

# 1 Introduction

At least since Herbert Simon's 1960 prediction that artificial intelligence would soon be able to replace all human labor, many economists have understood that sufficiently advanced artificial intelligence (AI) could transform the structure of the global economy. So far, of course, this has not occurred, but recent accelerations in the pace of AI development have renewed interest in AI's transformative potential. This review evaluates the channels through which sufficient advances in AI may affect economic growth by placing several strands of the literature on AI and growth within a common framework. We pay particular attention to implications for output growth, wage growth, and the labor share. For focus, we do not survey the other ways in which AI may have large economic impacts, such as by changing the income distribution.

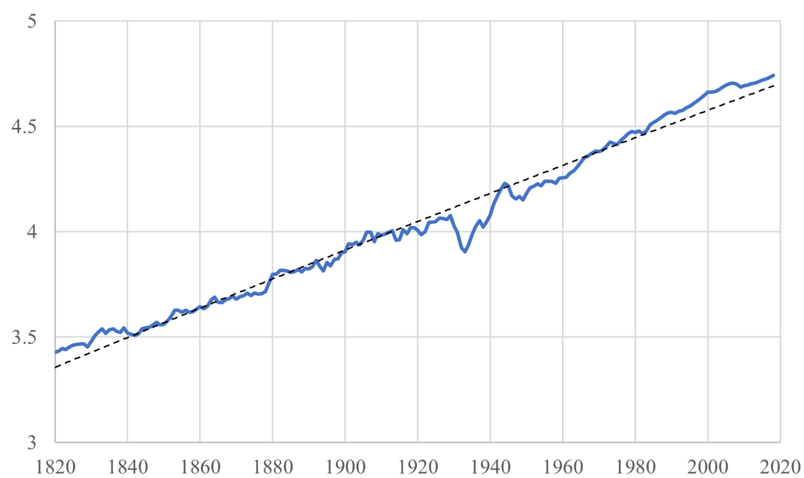


Figure 1: US gross domestic product per capita on a log scale since 1820, based on Maddison Project Database (Bolt and van Zanden, 2020) and US Bureau of Economic Analysis

For the past two centuries, economic growth in industrialized countries has largely conformed to two stylized facts first identified by Kaldor (1957). First, growth rates in output per capita have been approximately constant when averaged over long periods of time. Figure 1 illustrates this pattern in the context of the United States. Second, the fraction of output paid as wages—the labor share—has been approximately constant. Although Kaldor derived his stylized facts only by looking at time series from the United States and the United Kingdom from the mid-19th to the mid-20th century, they were soon shown to hold for other industrial countries and have held for the decades that followed. Much of the research on economic growth has since operated within frameworks that satisfy these stylized facts, which came to be cornerstones of economic growth theory for the Industrial Age. For a comprehensive overview of this body of research see, e.g., Acemoglu (2009).

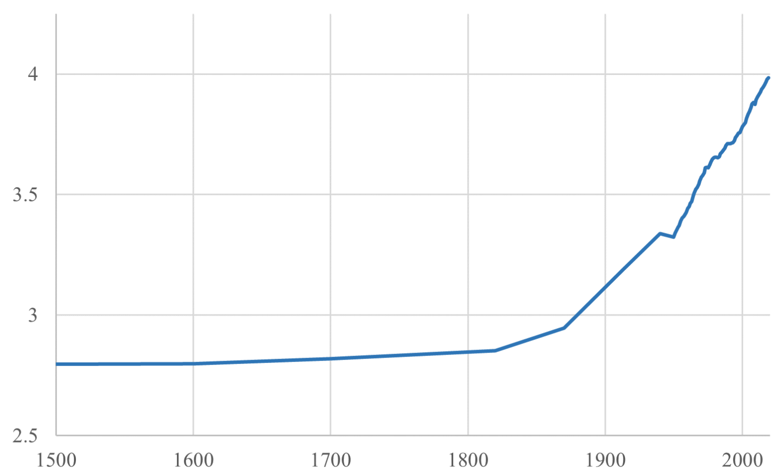


Figure 2: Gross world product per capita on a log scale since 1500. Data taken from Roodman (2020), Table 2

A different picture emerges when zooming out. The labor share in the distant past is difficult to estimate, and it does not appear have exhibited a clear long-run trend (Federico et al., 2020). But the economic growth rate has unambiguously increased over time when looking at sufficiently long time horizons and considering the world as a whole. Figure 2 depicts gross world output per capita on a log scale for the past five centuries. The phenomenon of constant growth since the Industrial Revolution now looks like a much more recent and gradual phenomenon, with the growth rate increasing as different countries transition to industrial growth at different times.[1]

We do not want to suggest that either of these figures is a reliable guide for predicting the exact path of economic growth in coming decades and

[1]Growiec (2022a) offers a complementary view of accelerating economic growth based on our growing capacity to process energy and information.

centuries. Instead, our objective in presenting Figures 1 and 2 is to highlight that near-constant exponential growth in output per capita is not a fact of nature. It has applied to specific countries over a specific period of history and may be a misleading guide to forecasting the future. In particular, if advances in AI prove equally or more transformative than the Industrial Revolution, the patterns of the Industrial Age that Kaldor described more than half a century ago may no longer hold.

Over the past decade, deep learning has made significant strides in performing individual cognitive tasks at human or super-human levels, from image and voice recognition to strategic games such as Go and scientific tools like AlphaFold. These systems are sometimes referred to as cases of "artificial narrow intelligence". While they have facilitated productivity growth, so far they have not fundamentally shifted economic growth patterns.

Current progress in the field of AI is driven by foundation models: large neural networks that are trained on vast amounts of data and then fine-tuned to perform an increasingly broad range of tasks (Bommasani et al., 2021). During the training process, these models develop what can be described as their own internal world models, on the basis of which they can generate outputs across a wide range of domains. The complexity of leading deep learning and foundation models has roughly quadrupled every year for the past decade (Sevilla et al., 2022) and has recently reached levels that are comparable to human brains (Carlsmith, 2020). Moreover, the capabilities of these models improve according to fairly predictable scaling laws: see Villalobos (2023) for a recent review.

The foundation models that have perhaps gained the most attention recently are large language models such as OpenAI's GPT-4, conveniently accessible in the form of the chatbot ChatGPT, or Google DeepMind's Gemini. These models can generate output that not only synthesizes existing knowledge but also develops new ideas based on the system's internal world model—in a manner similar to how humans create new ideas. Early analyses of the economic effects of even the current wave of generative AI suggest that it is likely to have a significant impact on growth, perhaps even doubling productivity growth over the next decade or more (Goldman Sachs, 2023; Baily et al., 2023).

Future models are likely to be capable of performing an even broader range of tasks and may eventually reach the level of Artificial General Intelligence (AGI), defined as the ability of AI systems to perform all cognitive tasks at least as well as a skilled human. In the process, unlike any other technology yet developed, such models could automate the process of research and development.

If AI capabilities cross certain thresholds, such as that of AGI, it is conceivable that a change in growth regimes as transformative as the Industrial Revolution would occur. Economic growth under the new regime would likely proceed more quickly than growth today. By changing the mechanics of growth, furthermore, such a transformation could also break the long-run constancy of the labor share: likely lowering it, and perhaps driving it to zero.

These effects are perhaps the intuitive ones. Hopes (and fears) of an AI-driven age of abundance without human work can be found throughout the popular imagination. Nevertheless, we consider it valuable to consider precisely whether and how such effects might unfold.

In recent years, economists have begun to engage earnestly in theoretical explorations of a wide array of the transformative possibilities of AI. We aim to synthesize the findings of these explorations.[2] In the process, we cover both the insights of different categories of models and the underlying mathematical intuition. We have simplified some of the models we cover and changed notation to fit them in a common framework and crystallize the mechanisms at work.[3]

The rest of this document proceeds as follows. Our basic setup and notation are introduced in the ensuing section, in the context of a minimal introduction to the economics of production, factor shares, and growth. Section 3 then discusses models in which AI is added to a standard production function. Section 4 discusses models in which AI is added to a "task-based" production function. Whereas both Sections 3 and 4 implicitly take place in settings of exogenous productivity growth, Section 5 discusses models in

[2] Sandberg (2013) presented an "overview of models of technological singularity" a decade ago, before mainstream economists started to analyze the transformative potential of AI. Most of the models he summarizes therefore do not attempt to spell out how AI or other transformative technologies would fit into standard economic models to produce the results in question. The models summarized here fill this gap.

[3] We do not offer all-things-considered assessments of the probabilities of particular AI-driven transformative scenarios. For one such assessment, see Erdil and Besiroglu (2023), who assign an approximately 50% chance to the event that, if AGI is developed this century, the economic growth rate will increase by at least a factor of ten.

which productivity growth is endogenous, with AI featuring in its development; and Section 6 places AI in final good production and technological development simultaneously. Section 7 compares the results found in Sections 3 to 6. Section 8 concludes.

# 2 Baseline model

## 2.1 Production and factor shares

Suppose that output depends on the input quantities of available production factors, and that we can categorize all production factors as either capital $K$ or labor $L$ and all output as a single good $Y$. Then we can capture production in the economy by a function $F(\cdot)$ that converts inputs into outputs as follows:

$$Y = F(K, L). \tag{1}$$

A production function $F(\cdot)$ will always be assumed to be continuously differentiable, increasing, and concave in each argument. It will also be assumed to exhibit constant returns to scale (CRS): doubling all the inputs to production would presumably double output.

The marginal product of a factor is the derivative of output with respect to that factor. In competitive markets, inputs are paid their marginal products. That is, the wage rate is $F_L(K, L)$, and the capital rental rate is $F_K(K, L)$, with the subscripts denoting partial derivatives. Intuitively, this reflects that in a competitive market with many firms facing a common production function, any factor that is being paid less than its marginal product will receive a higher offer by a competing firm, and any factor that is being paid more than its marginal product will be laid off.[4] The marginal productivity of each input is typically assumed to be increasing in the other input, i.e., $F_{LK} > 0$. Economically speaking, this captures that the inputs are complements.

[4]Most of our analysis assumes that markets clear, including that every worker is employed. This is an innocuous assumption if we are interested in a medium- or longer-term perspective of a world similar to what the US experienced over the past century (Figure 1), since unemployment fluctuated around a small and relatively constant natural rate of unemployment. However, if the world changes drastically and wages decline significantly, it may become important to also consider scenarios in which there is widespread unemployment (see e.g. Korinek and Juelfs, 2023).

Since $F(\cdot)$ is CRS, Euler's Homogeneous Function Theorem implies that the factor payments equal total output:

$$KF_K(K,L) + LF_L(K,L) = F(K,L). \tag{2}$$

This result implies that the output that is produced precisely covers what capital and labor earn, such that we do not need to be concerned about how to cover the factor payments or about what to do with any leftover output.

The fraction of output paid out as wages is termed the labor share:

$$s^L = LF_L(K,L)/F(K,L). \tag{3}$$

Likewise, the capital share $s^K$ is the fraction of output that accrues to the owners of capital. By (2), they sum to one: $s^K + s^L = 1$.

Throughout most of this review, we assume a production function that exhibits constant elasticity of substitution (CES).[5] Moreover, we assume that there are two technology parameters that reflect "capital-augmenting" technology, denoted $A$, and "labor-augmenting" technology, denoted $B$, which make the use of each respective factor more efficient, so that we can proceed as if we had more of it. In combination with the CRS assumption, these assumptions imply that, for some $\rho \leq 1$, production is given by

$$Y = [(AK)^\rho + (BL)^\rho]^{1/\rho} \tag{4}$$

if $\rho \neq 0$ and

$$Y = (AK)^a (BL)^{1-a}. \tag{5}$$

for some $a \in (0,1)$ for $\rho = 0$, i.e. the Cobb-Douglas case. The elasticity of substitution for this production function is $\sigma = 1/(1-\rho)$. (For readers unfamiliar with the concepts, Appendix A provides a formal definition of elasticity of substitution and additional properties of CES production functions.)

[5]In practice, of course, the substitution parameter between labor and capital may not be constant. It may be high when labor is more abundant than capital and low otherwise, or vice-versa. It may also change over time, for reasons independent of factor quantities. That said, it has been estimated in a variety of contexts and times to be substantially negative (see e.g. Oberfield and Raval (2021) or Chirinko and Mallick (2017)). When exploring preliminary hypotheses about growth, factor shares, and other macroeconomic variables, therefore, it can be helpful to start with a model in which production is CES (and CRS) with $\rho < 0$.

## 2.2 Exogenous growth

Over the past two centuries, as noted in the introduction, output per person and wages in the developed world have followed an exponential growth trajectory. In a two-factor production function *without* technology growth, the only possible explanation for growth in output per capita would be the capital stock growing more quickly than the population.

Capital accumulation cannot be the primary force driving long-run growth, however. This can most straightforwardly be seen in the context of CES production. If capital and labor are (weakly) gross complements ($\rho \leq 0$), then growth in the stock of capital per unit of labor reduces capital's marginal productivity to the point that growth in output per capita slows to zero. Analytically, we can see from equation (4) that the capital term tends to zero as the quantity of capital increases for $\rho < 0$. In the limit, then, if capital is far more plentiful than labor, output tends to $BL$, and output per person tends to $B$. Conversely, if capital and labor are gross substitutes ($\rho > 0$), capital accumulation can in principle sustain a positive growth rate.[6] However, in that case, the share of output earned by capital would converge to one, contradicting the Kaldor Fact of a roughly constant labor share.

Capital-augmenting technology growth cannot produce long-run growth in output and wages either—it just increases the effective capital stock $AK$, leading to the same results discussed in the previous paragraph.[7]

Uzawa (1961) showed that long-run growth in per-capita output and wages in this framework requires labor-augmenting technology growth. To illustrate this, suppose for simplicity that capital-augmenting technology $A$ and labor $L$ are fixed, that $B$ grows at some constant exponential rate $g_B$, and that a constant proportion $s > 0$ of output is saved as capital each period.[8] If the

[6] After accounting for capital depreciation, $\rho$ may have to be strictly positive for capital accumulation to allow long-run growth. We ignore capital depreciation throughout most of this document for simplicity.

[7] More generally, when highly complementary production factors undergo different rates of accumulation or factor-augmenting technology growth, the growth rate of output converges to the growth rate of the slowest-growing factor, and the share of that factor in output converges to 1. This is sometimes known as the "Baumol condition", after Baumol's (e.g. Baumol (1967)) seminal analyses of the increasing share of output spent on low-productivity-growth sectors, such as live entertainment.

[8] The saving rate is in fact historically (roughly) constant, at least in developed countries

saving rate $s$ is high enough that capital accumulation can keep up with the growing effective labor force, the result is a growth path in which output $Y_t$, capital $K_t$, and "effective labor" $B_tL$ all grow at rate $g_B$. By CRS, equal proportional increases to $K_t$ and $B_tL$ produce an equal proportional increase to $Y$.

To observe the importance of the parameter $s$, notice that if $s$ is too small, output is constrained by capital accumulation. (This is clearest when $s = 0$.) In this case, $Y_t$ eventually approximately equals $AK_t$. Then (using discrete-time notation for clarity),

$$K_{t+1} = K_t + sY_t \approx K_t + sAK_t, \tag{6}$$

so capital and output both grow at asymptotic rate $sA$.[9] The requirement that capital accumulation keep up with the growing effective labor force can be captured by the inequaliy $sA \geq g_B$. We call this the "sufficient saving" condition. Note that given any fixed $g_A > 0$ and $g_B$, the sufficient saving condition will eventually be met.

Letting the labor force grow at some positive rate $g_L$ makes no substantive difference. In this case, so long as $sA \geq g_B + g_L$, $BL$ grows at rate $g_B + g_L$, and $Y$ and $K$ do likewise. Regardless of population growth, factor shares are constant over time,[10] so wages and capital rents per person grow at rate $g_B$.

The empirical causes of technology growth remain highly uncertain. An exogenous growth model is one that does not attempt to model these causes, but simply takes constant exponential growth in $B$ as given.

# 3 AI in basic models of good production

## 3.1 AI and the productivity of capital

At face value, AI promises to make capital more productive. This would most naturally be modeled in the standard framework as an increase to $A$, which

over the past century or so. Section 3.5 provides microfoundations. For most of this document, we simply assume a constant saving rate.

[9]Models without labor, in which $Y_t = AK_t$, are termed "$AK$ models". An $AK$ economy is always constrained by capital accumulation and exhibits growth at rate $sA$.

[10]It follows from the identity $sY_t = K_{t+1} - K_t$ that $sY_t/K_t = g_{K,t}$. Since in the long run a constant saving rate maintains $g_Y = g_K$, the long-run capital share $(AK/Y)^\rho$ equals $(sA/g_Y)^\rho$. If there is insufficient saving (so that $g_Y = sA < g_B$), or in the edge case of $sA = g_B = g_Y$, the capital share tends to 1.

would amount to effective capital accumulation. As Acemoglu and Restrepo (2018a) point out, and as we have seen, this on its own would not be predicted to have very significant economic effects. It would increase output and wages somewhat. But given $\rho < 0$ and slow growth in labor supply, labor is the primary bottleneck to output, and any increases to wages would come ever more from an increase in the labor share rather than an increase in output. Indeed, the only "transformative" effect of capital-augmenting technology is that, as $A \to \infty$, all else equal, the labor share should rise to 1. This is of course the opposite of the intuitive trend, which is also the observed trend in the labor share in recent decades, especially in the industries that have undergone the most automation.[11]

The models below, therefore, are all designed to shed light on the consequences of increasing the productivity of capital *in combination* with various structural changes to the production function that AI might also precipitate.

## 3.2 Imperfect substitution

Nordhaus (2021) explores the transformative possibility of AI in the standard model of good production without adding anything explicit about AI. Instead, he posits that AI changes some of the model's parameters "behind the scenes", akin to an exogenous change in the parameters. This process has two steps.

First, suppose that AI raises the substitutability of labor and capital (or certain kinds of capital, such as computers) so that they become gross substitutes ($\rho$ is permanently bounded above 0). In this case, capital accumulation is sufficient for exponential output growth, even without population growth or technological development of any kind.

For illustration, consider our CES production function (4) with $\rho > 0$ and capital-augmenting technology $A$ as well as the saving rate $s$ held constant. If the capital supply grows more quickly than than the labor supply, $Y_t$ will come to approximately equal $AK_t$, and capital and output will accumulate exponentially at rate $sA$. More generally, if labor-augmenting technology grows exogenously at some rate $g_B \geq 0$, the output growth rate following the change in substitutability $\rho$ shifts from $\min(sA, g_B)$ to $\max(sA, g_B)$. The substitutability change thus increases the growth rate as long as $sA > g_B$.

Second, suppose that $A_t$ itself grows without bound. It does not matter

[11] See e.g. Acemoglu and Restrepo (2020, 2022) and Acemoglu et al. (2020).

whether this technology growth is due to AI or to forces that predated (but were less relevant before) the substitutability change. It also does not matter whether technology grows exogenously at some exponential rate $g_A > 0$, as Nordhaus (2021) assumes, or is the output of human research effort—in which case, even under constant population, $A_t$ rises without bound. In all cases, the growth rate of output will tend to $sA_t$, which, with $A_t$, will itself be growing indefinitely and without bound, delivering hyper-exponential growth. Aghion et al. (2019) call this type of growth acceleration whereby the growth rate itself increases without bound (but remains finite at any finite time horizon) a Type I growth explosion.

Under both transformative scenarios—the one-time growth rate increase that can occur without capital-augmenting technological development and the growth explosion that occurs with it—capital per worker will grow to infinity. However, since $\rho > 0$, the capital share will now tend to 1 rather than 0. For any fixed value $\rho < 1$, however, the marginal product of labor is still increasing in the supply of effective capital as $F_{LK} > 0$. Absolute wages will thus grow rapidly as the effective capital stock grows, as long as $\rho$ is bounded below 1. In fact, with $g_A > 0$, wages will grow superexponentially (though less quickly than output or effective capital). Absolute wages will stagnate only in the limit case that capital and labor become perfect substitutes ($\rho \to 1$), which we explore further in the ensuing section.

Nordhaus also discusses an analogous possibility: that AI may transform consumption growth by increasing the substitutability of goods on the demand side of the economy, rather than the substitutability of factors on the supply side. To explore this scenario, instead of dividing the space of goods into two production inputs and one output, let us divide it into one input ("capital" $K$) and two outputs (which might be called "standard consumption" $Y$ and "computer-produced consumption" $Z$). Capital grows exogenously at rate $g_K$. Given capital stock $K_t$, the production of the two consumption goods must satisfy

$$Y_t + Z_t/D_t = K_t. \tag{7}$$

That is, each unit of capital can produce either 1 unit of standard consumption or $D_t$ units of computer-produced consumption per unit time (without being used up). $1/D_t$ is the relative price of $Z$ at $t$: it is the number of units of $Y$ that must be given up at $t$ per unit of $Z$.

Consumers' utility functions all equal $U(\cdot)$, defined over $Y$ and $Z$. $U(\cdot)$ is homothetic, so consumers' elasticity of substitution between goods is everywhere defined, though not necessarily constant (see Appendix A). In response to consumer demand, production is allocated between $Y$ and $Z$ to maximize utility.

Suppose $D_t$ grows exponentially at rate $g_D$. This might be thought of as a direct implication of Moore's Law: famously, the number of computations that can be purchased with a given amount of capital seems to double approximately every eighteen to twenty-four months. The relative price of $Z$ then falls exponentially at rate $g_D$.

With each proportional fall in this relative price, the relative quantity of $Z$ demanded by consumers will rise by $\sigma_t g_D$, where $\sigma_t$ denotes the elasticity of substitution between the goods in the consumer utility function at $t$.

Now let $S_t$ denote the proportion of capital allocated to computing. The relative quantity of the computing good produced equals $Z_t/Y_t = D_t S_t/(1 - S_t)$. Considering the growth rate of this term, by the reasoning above, we have

$$g_D + g_{S,t} - g_{1-S,t} = \sigma_t g_D \tag{8}$$
$$\implies g_{S,t} - g_{1-S,t} = (\sigma_t - 1) g_D \tag{9}$$

If the two consumer goods are gross complements ($\sigma$ is bounded above by $\bar{\sigma} < 1$), this is always negative. Over the long run, $g_S$ must be negative and $g_{1-S}$ must be zero, since both terms are in the long run non-positive. Thus we have, in the long run, $g_S \leq (\bar{\sigma} - 1) g_D < 0$ and the fraction of capital employed to produce computer goods is ever-diminishing. Finally, since $Z = DSK$, we have

$$g_Z \leq \bar{\sigma} g_D + g_K. \tag{10}$$

In short, ever more computer goods are produced with an ever-declining fraction of the economy's capital.

On the other hand, if the two consumer goods are (weakly) gross substitutes ($\sigma_t$ is bounded strictly above 1), then the fraction of capital allocated to computing converges to one, so ultimately $g_Z = g_D + g_K$. (For $\sigma = 1$, the share of capital allocated to computing will be interior.)

The AI-relevant implications are straightforward. If computer-produced consumption is not currently very substitutable for other consumption ($\sigma$ bounded below 1), but developments in AI render it more substitutable (such that $\sigma$ is above 1), then the consumption growth rate could rise from something perhaps not much higher than $g_K$ to fully $g_D + g_K$. This would not be a growth explosion, as we are using the term. But given the speed of Moore's Law, it would have dramatic implications for consumer welfare.

Finally, Nordhaus constructs tests of the hypothesis that we are headed for a growth increase via the channels discussed above. If we are in fact headed for a supply-driven growth increase, we should expect to find a rising growth rate and a rising capital share. If we are headed for a demand-driven growth increase, we should expect to find a rising share of global income spent on computer-produced goods. He concludes that on balance the evidence does not support these hypotheses as of 2020.

Models in which labor and capital must be combined in more complex ways tend to produce the same broad conclusions. If labor and capital are sufficiently substitutable, then increasing capital-augmenting technology can increase the capital share, but it will still increase the absolute wage rate. Berg et al. (2018) detail a variety of such models.

## 3.3 Perfect substitution

We have seen that if labor and capital are the only two factors of production, then whenever the elasticity of substitution between them is finite, increases to the quantity of effective capital cause absolute wages to grow. Thus, if the elasticity shifts from less than 1 to greater than 1—a shift which can allow for faster capital accumulation—wage growth can accelerate, even as the labor share falls.

As we will see, however, prospects for wages look worse in cases of perfect substitutability. In this case, if there are only two production factors, the returns to each must be linear. Increases to the quantity of effective capital thus have no impact on the wage rate. If there are other production factors than "capital plus labor" (which are perfect substitutes), and if some become scarce relative to effective "capital plus labor", increases to the quantity of effective "capital plus labor"—driven by increases in effective capital—drive wages down.

A model of the beginning of perfect substitution between labor and capital can be presented most straightforwardly as one in which human-substitute robots are simply at first expensive, and then cheap, in units of human labor hours. This is because, when goods are perfect substitutes toward some end, they are only ever both purchased in positive quantities when their prices are the same. Even if it were already feasible to produce robots fully substitutable for human labor, therefore, we would only see any produced, and observe their effects, once their rental rate had fallen below what would otherwise have been the wage rate. In other words, perhaps the substitutability does not need to rise; perhaps it is perfect, and all that needs to change is a relative price.

To illustrate this dynamic, consider the following simple model, inspired by Hanson (2001). Equipment $Q$, labor $L$, and land $W$ are employed in a Cobb-Douglas production function,

$$Y = F(Q, L, W) = Q^a L^b W^{1-a-b}. \tag{11}$$

Since the quantity of land is fixed, we can normalize it to 1 and drop it from the expression, remembering that $a + b < 1$.

The output good can be consumed or invested as capital $K$. Capital can serve either as equipment or as robotics, which functions as labor, whereas the human workforce $H$ is fixed and can only serve as labor. The level of capital-augmenting technology is denoted $A$. That is, if $S$ denotes the fraction of capital employed as robotics, output is

$$Y = ((1 - S)AK)^a (H + SAK)^b. \tag{12}$$

$A_t$ rises exogenously without bound. For simplicity we will assume that a constant and sufficient fraction $s$ of output is saved as capital. Because the substitution parameter between equipment and labor is not less than (in fact is equal to) 0, the accumulation of effective equipment is enough to sustain output growth.

Early in time, when effective capital is scarce, all capital is used as equipment; $S = 0$. Indeed, at the rate at which capital can be converted from equipment to robotics, it would be valuable instead to use some human labor as equipment, if that were possible. Capital then grows (using discrete-time notation for clarity) such that

$$\begin{aligned} K_{t+1} &= K_t + s(A_t K_t)^a H^b \qquad (13)\\ \Rightarrow \quad g_{K,t} &= (K_{t+1} - K_t)/K_t = s A_t^a K_t^{a-1} H^b. \end{aligned}$$

As we can see from the right hand side, capital growth will approach a steady state such that

$$a g_A + (a - 1) g_K = 0 \quad \Rightarrow \quad g_K = \frac{a}{1 - a} g_A. \tag{14}$$

We will thus have output growth of $g_Y = a(g_A + g_K) = g_A a/(1-a)$.

As the equipment stock grows, wages rise. As capital-augmenting technology rises and effective equipment grows more abundant, however, there comes a time past which it is optimal to split further capital between equipment and robotics. The labor growth rate then jumps to the rate that keeps its marginal productivity equal to that of equipment, and the output growth rate jumps accordingly.[12] In particular, with capital now filling the roles of both equipment and labor, we now have

$$g_Y = g_K = \frac{a+b}{1-a-b}\, g_A, \tag{15}$$

by the same calculation as above.[13]

Hanson estimates the growth implications of crossing the robotics cost threshold using a slightly more realistic model with roughly realistic estimates of the parameters involved. The level of capital-augmenting technology is assumed to double (i.e. the cost of effective capital is assumed to halve) every two years, in a conservative approximation to Moore's Law. Before capital begins to be used as robotics, output in the model grows at a relatively familiar rate of 4.3% per year. After, the growth rate is 45%.

In the model above, because the production function is Cobb-Douglas, the labor share—the share of output paid in compensation for human and/or robotic labor—is independent of the factor quantities. As human labor constitutes an ever smaller share of total labor, however, the *human* labor share falls to zero.

Furthermore, even the absolute wage $F_L$ falls to zero. To see this, note that in a CRS production function, the marginal productivities of equipment and labor are kept equal ($F_{Q,t} = F_{L,t}$) when the quantities of the two factors grow at the same rate. We can thus rearrange our formula regarding

[12] As Yudkowsky (2013) points out, we might interpret this as a model in which AI comes in the form of " emulations"—a theoretical technology on which Hanson has written extensively—which are always technically feasible but which are, at first, prohibitively expensive, because effective capital is sufficiently scarce.

[13] Note that, were it not for the inclusion of the non-accumulable factor land, there would be no steady-state growth rate; in solving for it, we would have to divide by 0. Instead, the economy would be, asymptotically, an $AK$ economy with exogenous capital-augmenting technology growth. As we saw in the previous section, we would have a Type I growth explosion.

competitive CRS factor payments:

$$F_{L,t}Q_t + F_{L,t}L_t + F_{W,t} = Y_t \;\Rightarrow\; F_{L,t} = \frac{Y_t - F_{W,t}}{Q_t + L_t}. \tag{16}$$

With a constant share of output accruing to land as well, but the quantity of land fixed, land rent per unit of land—i.e. the land rental rate $F_W$—must grow at the same rate as output. $Y$ and $F_W$ will thus both grow at $g_Y$, and $Q$ and $L$ will both grow at rate $g_A + g_K = g_A + g_Y > g_Y$. The right-hand ratio will then fall to zero.

In any CRS production function without labor-augmenting technology, what happens to the marginal productivity of labor, and thus wages, depends on the quantity of effective labor relative to that of the other effective factors of production. This relative quantity need not rise; it could fall if labor's complements grow productive and plentiful more quickly than its substitutes, or stay fixed if they grow at the same rate.

Consider the following model, very similar to the above, but in which technology augments only equipment, not capital used as robotics:

$$Y = F(Q, L[, W]) = Q^a L^b = ((1-S)AK)^a (H + SK)^b. \tag{17}$$

The growing stock of equipment implies that, as above, wages rise before the substitutability cost threshold is crossed. Furthermore, we will still have $g_Y = g_K$, and thus

$$g_Y = a(g_A + g_K) \;\Rightarrow\; g_Y = \frac{a}{1-a}\, g_A \tag{18}$$

before the threshold is crossed. Finally, the threshold will still eventually be crossed: if all capital were used as equipment indefinitely, the marginal productivity of capital used as equipment $AF_Q = aY/K$ would fall below that of labor $F_L = bY/H$.

After the threshold is crossed, we will have

$$g_Y = a(g_A + g_K) + b g_K \;\Rightarrow\; g_Y = \frac{a}{1-a-b}\, g_A. \tag{19}$$

This is still a growth rate increase, though not as large as that in the first model above.

To see what happens to wages, however, observe that when $S$ is chosen so that invested output is split optimally between equipment and robotics, it will satisfy $AF_Q = F_L$, or

$$\frac{aY}{(1-S)K} = \frac{bY}{H+SK} \quad \Rightarrow \quad S = \frac{b - aH/K}{a+b}. \tag{20}$$

As $K \to \infty$, we have $S \to b/(a+b) < 1$, and therefore $g_L = g_K = g_Y$. As above, because the production function is Cobb-Douglas, the labor share is constant. Now, however, the quantity of effective labor grows no more quickly than output, so labor payments per labor quantity—i.e. wages per human worker—merely stagnate.

Korinek and Stiglitz (2019) offer another illustration of this phenomenon, in the context of a somewhat similar model. As usual we will simplify here to highlight the intuition.

Suppose that $Y$ is produced as in the second model of this section (i.e. the model just above), except that the substitution parameter between land and the other two factors is bounded below 0. Though land is in fixed supply, it is at first plentiful enough that its factor share is low. The saving rate is fixed.

At first, as capital accumulates, it is split between use as robotic labor and use as equipment, so that the relative quantities of labor and equipment are unchanged. The capital and labor shares are roughly constant, but the absolute wage stagnates, as we have seen. In time, however, land becomes a binding constraint. The share of output received as land rents approaches 1, and the absolute wage falls to 0.

As should be clear, the same logic could apply to many more complex models. Embed *any* production function in a "surrounding" production function with a fixed-supply and low-substitutability resource such as land, and in the long run, even if all of the original production function's resources grow abundant, the resource in fixed supply constrains growth and its owners receive approximately all output.

## 3.4 Substitutability in robotics production

Like Korinek and Stiglitz, Mookherjee and Ray (2017) develop a model in which capital can replace human labor without technological progress. Unlike Korinek and Stiglitz, they do not simply assume that capital can be used as robotics, but make the robot production function explicit and identify a condition under which human labor replacement can occur. A simplification of their model is as follows.

The final good $Y$ is produced using capital $K$ and labor $L$ in a typical two-factor production function $F(\cdot)$, with a substitution parameter bounded below 0. Labor is supplied by human work $H$ and robotics $R$, which are perfect substitutes. Robotics is better thought of as the provision of robot *services* than as robots, because it must be used as it is produced; it cannot be accumulated. If we would like to think of it as a kind of physical capital, we would say that it exhibits full depreciation.

Robotics is also produced using capital and labor, using a standard but perhaps different production function $f(\cdot)$, also with a substitution parameter bounded below 0. Whereas one unit of robotics is defined as that which replaces 1 human worker in the output production function, one unit of robotics replaces some $D \in (0,1)$ human workers in the robotics production function. For each input $X \in \{K, H, R\}$, $S_X$ is defined (assuming $X > 0$) to be the fraction of $X$ that is allocated to the production of robotics rather than the final good. For simplicity, the population of human workers is fixed and there is no technological progress. More formally, output and robotics at $t$ are

$$\begin{aligned} Y_t &= F((1-S_{K,t})K_t, (1-S_{H,t})H + (1-S_{R,t})R_t); \\ R_t &= f(S_{K,t}K_t, S_{H,t}H + DS_{R,t}R_t). \end{aligned} \tag{21}$$

As usual, a constant fraction of output is saved as capital.

Early on, when capital is scarce and human labor relatively plentiful, there may be no reason to produce robotics at all. As capital accumulates and output begins to be constrained by human labor, however, the marginal output productivity of capital falls to zero. It may therefore at some point be worthwhile to allocate some positive fractions $S_K$ and $S_H$ of available capital and human labor to robotics production. To be precise, it will necessarily start being worthwhile iff $f_L(k,0) > 1$ given $k > 0$, i.e. if, given some capital (which is eventually near-worthless in final good production), marginal contributions of labor can create robotics at a ratio of more than 1:1. Let us call this the "robotization condition". Note that it is a relatively weak condition; $f_L(k,0) = \infty$ given $k > 0$ if $f(\cdot)$ is CES, for example.

If robotics production relies on capital and *human* labor—i.e. if we set $S_R = 0$—it too will ultimately be constrained by lack of labor:

$$\text{as } S_{K,t}K_t \to \infty,\ R_t \to S_{H,t}\bar{R}, \text{ where } \bar{R} \triangleq \lim_{x\to\infty} f(x,H). \tag{22}$$

Output in turn is constrained by total labor, despite the possibility of robotics:

$$\text{as } (1 - S_{K,t})K_t \to \infty,\ Y_t \to \bar{Y}\frac{(1 - S_{H,t})H + R_t}{H + \bar{R}}, \tag{23}$$
$$\text{where } \bar{Y} \triangleq \lim_{x\to\infty} F(x, H + \bar{R}).$$

Long-run output will be maximized by setting $S_H = S_H^*$, the value that maximizes $(1 - S_H^*)H + S_H^*\bar{R}$—i.e. $S_H^* = 1$ if the robotization condition is met, because in this case $\bar{R} > H$, and $S_H^* = 0$ if not. Long-run output will then approach an upper bound of $\bar{Y}((1-S_H^*)H+S_H^*\bar{R})/(H+\bar{R})$. The human labor share will approach 1, either because it is the scarce input to output directly or because robotics is the scarce input to output and human labor is the scarce input to robotics. The absolute wage will of course stagnate. In short, robotization can raise the output ceiling, but it cannot on its own produce a sustainably positive growth rate.

One might expect that, if we do not fix $S_R = 0$, it will eventually be optimal to use robotics in the production of robotics. In fact, this will only happen if $D$ is large enough that, as the quantity of capital allocated to robotics production grows large, one unit of robotics can produce more than one unit of robotics: that is, if $\lim_{k\to\infty} f_L(k, H) > 1/D$; or equivalently, because $f(\cdot)$ is CRS, if $f_L(k, 0) > 1/D$ for $k > 0$. The identification of such a condition in their more general setting is Mookherjee and Ray's key insight, and they call the condition the "von Neumann singularity condition", after the work by Burks and Von Neumann (1966) on self-replicating automata. It is of course very closely analogous to the robotization condition above, but stronger since we are imposing $D < 1$. We might take this to be the natural case; robotics production is presumably harder to automate than most other tasks are.[14]

Suppose that this condition is met, and that $S_K K$ is large enough that $f_L(S_K K, H) > 1/D$. Then there is an optimal quantity of robotics to allocate to robotics production, so as to maximize *net* robotics production. This is the quantity such that use of a marginal unit of robotics on robotics production increases robotics output by exactly one unit. That is, it is optimal to set $S_R > 0$ such that $S_R$ satisfies $f_L(S_K K, H + DS_R R) = 1/D$.

[14]Presumably at least some other tasks are more difficult to automate, however. As Mookherjee and Ray present in the original paper, their central result does not depend on robotics production being more difficult to automate than *all* other tasks, just on robotics production being sufficiently difficult to automate.

The value of $R$ depends on $S_R$, since more inputs to robotics production will correspond to higher robotic output. Nevertheless, we know that a unique $S_R \in (0,1)$ satisfying the above equality exists, for a given $S_K K$. To see this, recall that $f_L(S_K K, H + DS_R R) > 1/D$ at $S_R = 0$, by supposition. And we must have $\lim_{S_R \to 1} f_L(S_K K, H + DS_R R) < 1/D$, or else the quantity of robotics output $R$ and thus also $S_R R$ would grow without bound, fixing $S_K K$, as $S_R \to 1$; but in this case $f_L \to 0$, by the assumption that the substitution parameter in the robotics production function is bounded below 0. By the concavity and continuous differentiability of $f(\cdot)$, therefore, there is a unique $S_R^* : f_L(S_K K, H + DS_R^* R) = 1/D$.

Under the singularity condition, $S_K$ and $S_R$ approach constants $S_K^*$ and $S_R^*$, strictly between 0 and 1, as the capital stock grows.[15] Growth proceeds as in an $AK$ model, with the rate of capital accumulation, final good output growth, and robotics output growth all asymptotically constant and proportional to the saving rate.[16] The wage level is constant and lower than it is in the absence of the singularity condition, since the ratio of capital to labor in robotics production is still asymptotically constant but now positive rather than zero. The share of income accruing to human labor falls to zero.

As with Hanson (2001), moderate tweaks to this model could result in absolute human wages rising or falling, rather than merely stagnating. Also, in the presence of population growth or labor-augmenting technology growth, introducing automation can increase the growth rate of final good output (or final good output per capita) from the rate of effective labor (or labor-augmenting technology) growth to something much higher, given sufficient saving.

As we saw in Section 3.3, when human work must compete with robotics for which it is perfectly substitutable, the standard result is that the human labor share falls to 0 and the wage stagnates or changes (likely falls) as well. Above, however, we saw that modeling robotics production explicitly, rather than stipulating a frictionless conversion of the final output good into robotics, allows for a channel through which human work can remain necessary. A

[15] See Appendix B.1 for a proof.

[16] The model will approximate an $AK$ model with $A = Y/K = F((1-S_K^*), (1-S_R^*)R/K)$, where $R/K$ likewise satisfies $R/K = f(S_K^*, S_R^* R/K)$. Technological progress that allows capital to produce more robotics increases long-run $R/K$ and functionally "increases $A$", though it will never exceed its upper bound of $F(1, \infty)$. This will be finite, by the assumption that $F(\cdot)$'s substitution parameter is bounded below 0.

positive human labor share can be maintained, even when robotics can fully substitute for human work in the final good production function, when human work cannot be fully substituted for in *robotics* production.

Korinek (2018) presents another model in which robots are not simply "capital that can function as labor" but are economic agents of their own that must be sustained through robot absorption, just like humans need consumption to be sustained. He too finds that the human labor share, and in his case also the wage rate, can fall to 0 unless human labor remains necessary for the maintenance of the robot population.

Growiec et al. (2023) develop a framework in which they distinguish production factors into "hardware" and "software" rather than the typical factors capital and labor. They observe that any production process requires physical action steered by information processing. Traditionally, these two functions have been performed by human brawn and human brains, but technological progress over the centuries has increasingly enabled machines to do both.

## 3.5 Growth impacts via impacts on saving

In some of the models we have considered, saving has been an important determinant of growth. The saving rate, however, has been assumed to be exogenous. This leaves open another channel through which developments in AI could impact growth: by changing the rate of return to saving, more advanced AI could change the rate of saving and thus the growth rate.

This scenario can be illustrated most simply using a model from Korinek and Stiglitz (2019). Suppose labor and capital are perfectly substitutable.[17] Labor can only be supplied by humans. Activity unfolds in discrete time, and capital depreciates fully every period; it cannot accumulate. (Capital depreciation simplifies the exposition but is not necessary for the central result.) Given saving rate $s_t$, output and capital growth are then

$$Y_t = AK_t + BL_t, \quad K_{t+1} = s_t Y_t. \tag{24}$$

We begin with $K = 0$. Output per capita is thus $B$ and, without saving, does not grow. If $A < 1$, there is no incentive to save, and doing so cannot generate growth; foregoing each unit of consumption at $t$ would offer someone

[17] Equivalently, we could say that output is produced by a single factor, labor, which can be supplied both by humans and by robots.

only $A < 1$ additional units of consumption at $t + 1$, starting from the same baseline of $B$. For any $A > 1$, on the other hand, sufficiently patient individuals will want to save some fraction $s_t > 0$ of their incomes; not to do so would be to miss the opportunity to give up marginal consumption at baseline $B$ in exchange for a larger quantity of marginal consumption also at baseline $B$. More precisely, it should be clear that positive saving will be optimal, given any pure time discount factor $\beta < 1$, as long as $A > 1/\beta$. Furthermore, under certain assumptions about the shape of individuals' utility functions, the induced saving rate will be some constant $s > 1/A$, independent (at least roughly) of the absolute output level. In the long run, as the relative contribution of effective human labor $BL_t$ grows negligible, we will have $Y_t \approx AK_t$. And since

$$\begin{aligned} K_{t+1} = sY_t \approx sAK_t \\ \Rightarrow (K_{t+1} - K_t)/K_t \approx sA - 1, \end{aligned} \tag{25}$$

capital and therefore output will grow at asymptotic rate $sA - 1 > 0$.

In short, an increase in $A$—induced, perhaps, by AI developments which render robots cost-effective replacements for human labor—can trigger saving and can thus increase the growth rate of output and output per capita. Here, an $A$-increase raises per capita output growth from zero to a positive number, leaves the wage rate constant at $B$, and pushes the human labor share to zero; but other impacts on wages, the labor share, and growth are possible. The point is just that, in addition to the ways in which increases in $A$ can sometimes directly impact the growth rate, they can sometimes do so indirectly by impacting $s$.

There is another mechanism through which developments in AI could impact the saving rate. If the saving rate is heterogeneous across the population, then growth will depend on how income is distributed between high- and low-savers. Developments in AI could thus affect the growth rate by affecting the income distribution. In principle, this effect could have implications for growth in either direction. Here, we will focus on the especially interesting and counterintuitive possibility that AI slows and even reverses growth by transferring wealth from those with low to those with high propensity to consume.

This scenario is illustrated most simply by Sachs and Kotlikoff (2012), though the same mechanism is explored in more detail by Sachs et al. (2015).

If some investment goods are sufficiently substitutable for labor, automation raises capital rents but lowers wages. If saving for the future comes disproportionately out of wage income, for whatever reason, then this wage-lowering can cause future output to fall.

Consider an overlapping generations economy with constant population size. Each person lives for two periods. The young work, investing some of their income; the old live off their investments. More precisely, output is a symmetric Cobb-Douglas function of labor and capital. The output good can be consumed or invested as capital $K$. Capital can be used either as equipment $Q$ or as robotics, which serves as labor, and it is split between these uses until their marginal products are equal. The human workforce $H$ is fixed and can only serve as labor. Unlike in the Hanson model, however, equipment's factor-augmenting technology level is fixed. $A$ denotes the augmentation of robotics only.

Formally, if $S$ is the share of capital used in robotics,

$$Y_t = F(Q_t, L_t) = ((1 - S_t)K_t)^{1/2}(H + S_t A_t K_t)^{1/2}. \tag{26}$$

Capital at $t$ is financed by those who were young at $t-1$, who put aside half their wage incomes as investment.[18] The old at $t$ consume all their wealth: not only their investment income, $F_{Q,t}(1 - S_t)K_t + F_{L,t}S_t A_t K_t$, but even the capital stock $K_t$, which is liquidated after use in production. The economy is in a zero-growth steady state when investment is just replenished each period: that is, when $F_{L,t}/2 = K_t$.

Now suppose robotics grows slightly more productive, so that $A_{t+1} > A_t$. Let $G_A \triangleq 1 + g_A$ denote $A_{t+1}/A_t$. For a single period, total output and the incomes of the old grow. The young see a fall in wages, however, and investment therefore falls as well. This fall in investment outweighs the fact that some of the investment, namely that in robotics, is now more productive. Output therefore falls. The wage rate falls too, due both to the abundance of robotics and to the lack of investment in equipment.

More formally: in the new equilibrium, the marginal product of robotics is again equal to that of equipment. Because the "relative cost" of robotics has now been multiplied by $1/G_A$ and because here $\rho = 0$, the relative quantity of

[18] Let $r_t$ denote the interest rate at $t$: that is, here, $F_{Q,t+1}$, or equally $A_{t+1}F_{L,t+1}$. Suppose that period utility is logarithmic in consumption and that the young choose the saving rate $s_t$ to maximize lifetime utility $\ln((1 - s_t)F_{L,t}) + \ln(s_t F_{L,t}(1 + r_t))$. Then the chosen saving rate will always equal 1/2.

labor must be $G_A$ times higher. Letting asterisks denote the new equilibrium outcomes, $L^*/Q^* = G_A L_t/Q_t$. So

$$\begin{aligned} F_L^* &= \frac{1}{2}\Big(\frac{L^*}{Q^*}\Big)^{-1/2},\ F_{L,t} = \frac{1}{2}\Big(\frac{L_t}{Q_t}\Big)^{-1/2} \\ \Rightarrow F_L^* &= F_{L,t} G_A^{-1/2} < F_{L,t}. \end{aligned} \tag{27}$$

The new rate of return on investment is $r^* = A_{t+1}F_L^* = G_A A_t F_L^*$, the new wage is $F_L^*$, and the saving rate remains 1/2. The consumption of the old thus equals half their income while young, times $1 + r^*$:

$$\begin{aligned} (1 + G_A A_t F_L^*)\frac{F_L^*}{2} &= (1 + A_t G_A^{1/2} F_{L,t})\frac{F_{L,t}}{2G_A^{1/2}} \\ = (G_A^{-1/2} + A_t F_{L,t})\frac{F_{L,t}}{2} &< (1 + A_t F_{L,t})\frac{F_{L,t}}{2}. \end{aligned} \tag{28}$$

The new equilibrium therefore features lower output in all subsequent periods, and lower consumption for both young and old.

If robotics-augmenting technology continues to grow at rate $g_A$, the wage rate, and thus the consumption of the young, will continue to fall to 0 at rate $G_A^{1/2} - 1$. The consumption of the old (and thus also output) will fall at a falling rate, and will eventually stabilize above 0, as the increasing productivity of invested equipment ever more closely compensates for the falling absolute amount invested.[19] The human labor share thus falls to 0.

Again, the direction of these impacts is sensitive to whether the "winners" from advances in AI save more or less than the "losers". As Berg et al. (2018) point out, for instance, those who make most of their incomes from wages currently empirically exhibit lower saving rates than those who make most of their incomes from capital rents, so the mechanism identified by Sachs and Kotlikoff should if anything increase output growth. In any event, the key point is essentially a special case of the well-known fact that, in a neoclassical growth model with finitely lived agents and no (or imperfect) intergenerational altruism, the rate of saving is not necessarily optimal. Accordingly, policymakers must always consider not only the impact of a policy or technological development on short-term output, but also its impact on the saving rate. When a given development produces a suboptimal saving rate, it should

[19] Technically, if flow utility is logarithmic in consumption (see footnote 18), lifetime utility falls to negative infinity as the consumption of the young falls to zero.

be counterbalanced by investment subsidies or by transfers from those with high to those with low propensity to consume: in this model, from the old to the young.

# 4 AI in task-based models of good production

## 4.1 Introducing the task-based framework

In Section 3, we imagined that capital and labor were each employed in a single sector. In the Cobb-Douglas case, we held the exponent $a$ on capital fixed. We then explored the implications of changing $\rho$, the substitutability of capital and other durable investments for human labor, and of independently changing the growth rates of factor-augmenting technology.

In reality, however, capital and labor are of course employed heterogeneously, and this heterogeneity seems likely to shape the economic impacts of developments in AI. Indeed, sectors with high rates of automation have historically experienced stagnating or declining wages (Acemoglu and Autor, 2012; Acemoglu and Restrepo, 2020), even as wages on average have increased.

Here, therefore, we will explore a model of CES automation from Zeira (1998), which makes room for this sort of heterogeneity. (We will follow the exposition and extension of Zeira's model given by Aghion et al. (2019).) As we will see, this model amounts roughly to assuming a *fixed* substitution parameter $\rho$ and either a changing capital exponent $a$, in the Cobb-Douglas case, or impacts on factor-augmenting technology which are sensitive to $\rho$ when $\rho \neq 0$.

Let us begin with the $\rho = 0$ case. Suppose output is given by a Cobb-Douglas combination of a large number $n$ of factors $X_i$, for $i = 1, \ldots, n$:

$$Y = X_1^{a_1} \cdot X_2^{a_2} \cdot \cdots \cdot X_n^{1-a_1-\cdots-a_{n-1}}. \tag{29}$$

At such a fine-grained level, these "factors" might better be thought of as intermediate production goods (Zeira, 1998), or even as individual tasks (Acemoglu and Autor, 2011). We will refer to them as tasks.

Fraction $a$ of the tasks are automatable, in that they can be performed by capital or labor, and fraction $1-a$ are not, in that they can only be performed by labor. Given capital and labor stocks $K$ and $L$, if all automatable tasks are indeed automated (performed exclusively by capital), $K/(na)$ units of capital will be spent on each automatable task and $L/(n(1-a))$ units of labor on each non-automated task. With just a little algebra, we have

$$Y = AK^a L^{1-a}, \tag{30}$$

a two-factor Cobb-Douglas production function with an unimportant coefficient $A$.[20]

Now consider a CES production function, of substitution parameter $\rho \neq 0$, with a continuum of production factors $Y_i$ from $i = 0$ to 1, instead of just two:

$$Y = \Big( \int_0^1 Y_i^\rho di \Big)^{1/\rho} \tag{31}$$

Tasks $i \leq \beta \in (0, 1)$ are automatable.

Let $K$ and $L$ denote the total supplies of capital and labor, and $K_i$ and $L_i$ the densities of capital and labor allocated to performing some task $i$ (so $Y_i = K_i + L_i$). Suppose again that all automated tasks are indeed performed exclusively by capital. Since the tasks are symmetric and the marginal product of each task is diminishing ($\partial^2 Y/\partial X_i^2 < 0$), the density of capital applied to each task $i \leq \beta$ will be equal (assuming that production proceeds efficiently), as will the density of labor applied to each $i > \beta$. And since we must have

$$\int_0^\beta K_i di = K \text{ and } \int_\beta^1 L_i di = L, \tag{32}$$

we know that $K_i = K/\beta \ \forall i \leq \beta$ and $L_i = L/(1-\beta) \ \forall i > \beta$. We can thus write our production function as

$$\begin{aligned} Y &= [\beta(K/\beta)^\rho + (1-\beta)(L/(1-\beta))^\rho]^{1/\rho} \\ &= [\beta^{1-\rho} K^\rho + (1-\beta)^{1-\rho} L^\rho]^{1/\rho} \\ &= F(AK, BL) = [(AK)^\rho + (BL)^\rho]^{1/\rho} \end{aligned} \tag{33}$$

where $A = \beta^{(1-\rho)/\rho}$ and $B = (1-\beta)^{(1-\rho)/\rho}$. This is simply a two-factor CES production function.

[20] $A = a^{-a}(1\text{–}a)^{a-1}$, which ranges from 1 (at $a = 0$ or 1) to 2 (at $a = 1/2$).

We have assumed that all automatable tasks are indeed performed exclusively by capital. This will obtain so long as there is more capital per automatable task than labor per non-automatable task, i.e. as long as

$$\frac{K}{\beta} > \frac{L}{1-\beta}. \tag{34}$$

In this case, even when capital is spread across all automatable tasks, we have $F_K < F_L$, so there is no incentive to use labor on a task that capital can perform. Let us call this condition the "automation condition". For any fixed $\beta$, if the capital stock grows more quickly than the labor supply, the condition will eventually hold.

## 4.2 Task automation

Let us now explore the implications of task automation in more detail, across the regimes of CES production with $\rho = 0$ and $< 0$. The case of $\rho > 0$ will be covered in Section 4.3.

As we have seen, under Cobb-Douglas production, task automation raises $a$ along the range from 0 to 1. The Cobb-Douglas production function under factor-augmenting technology is given by (5). Since a constant saving rate imposes $g_Y = g_K$ in the long run, the long-run growth rate in this case satisfies

$$\begin{aligned} & g_Y = a(g_A + g_K) + (1-a)(g_B + g_L) \\ \Rightarrow \quad & g_Y = \frac{a}{1-a} g_A + g_B + g_L. \end{aligned} \tag{35}$$

The impact of a one-time increase to $a$, or of increases only up to some bound strictly below 1, is therefore straightforward. The capital share rises with $a$. If $g_A > 0$, the growth rate increases, ultimately raising the wage rate; otherwise the growth rate is unchanged, and the impact on the wage rate is ambiguous.

Given asymptotic complete automation, with $1 - a$ falling to 0 exponentially or faster, the model approximates an $AK$ model. If $g_A = 0$, the growth rate rises to $sA$. If $g_A > 0$, the growth rate rises without bound, and wages too rise superexponentially.

Automation, as we have defined it, allows capital to perform more tasks. One might therefore imagine that it is equivalent to the development of some sort of capital-augmenting technology. Aghion et al. (2019) observe, however, that automation in the above model is actually equivalent to the development of *labor*-augmenting—and capital-depleting!—technology, as long as $\rho < 0$ and the automation condition holds. To see this, recall our production function:

$$Y = [(AK)^{\rho} + (BL)^{\rho}]^{1/\rho}, \tag{36}$$

where $A = \beta^{(1-\rho)/\rho}$ and $B = (1-\beta)^{(1-\rho)/\rho}$. As $\beta$ rises from 0 to 1, therefore, $A$ falls from unboundedly large values to 1, and $B$ in turn rises from 1 without bound.

The reason for this result is that, as $\beta$ rises, capital is spread more thinly across the widened range of automatable tasks, and labor is concentrated more heavily in the narrowed range of non-automatable tasks.[21] Automation therefore allows capital to serve as a *better complement* to labor. A marginal unit of labor is spread across fewer non-automatable tasks, producing a larger increase to the supply of each; given the abundance of capital, this then produces a larger increase to output. Conversely, under this allocation, labor serves as a worse complement to capital, requiring capital to spread itself over more tasks (and only partially compensating for this effect by supplying the remaining tasks more extensively).[22]

As explained in Section 2.2, when $\rho < 0$, labor-augmenting technology is the key to sustained output growth. Let us spell that out in this context. Suppose that, by some exogenous process, a constant fraction of the remaining non-automatable tasks are made automatable each period, so that $(1-\beta_t) \to 0$ at a constant rate $g_{1-\beta} < 0$. Then $B$ will grow at rate $g_B = g_{1-\beta}(1-\rho)/\rho > 0$. $A$ is asymptotically constant at 1, so $g_A \approx 0$. If the saving rate $s$ is constant and high enough to maintain the automation condition, we get our familiar "balanced growth path". The capital stock, effective labor supply, and output all grow at asymptotic rate $g_Y = g_B + g_L$, output per capita grows at rate $g_B$, and the labor share is asymptotically

[21] Here we are only considering increases in $\beta$ up to the point that capital per automatable task no longer exceeds labor per non-automatable task, so that the automation condition is satisfied.

[22] These observations are also made by Growiec (2022b) who shows that under partial automation with $\rho < 0$, capital and labor remain complements. However, under full automation, they become substitutes.

constant and positive.[23]

Automation can thus increase the growth rate of output per capita, and have other transformative consequences.[24] In the model above, because the automation rate $-g_{1-\beta}$ is the *only* driver of growth, introducing it increases the growth rate from 0 to $g_{1-\beta}(1-\rho)/\rho$. In the presence of growth from other sources, automation can increase the growth rate further. Consider for instance what follows if we have $B_t = D_t(1-\beta)^{(1-\rho)/\rho}$, with $\beta$ constant but $D_t$ growing exogenously at rate $g_D$. Given saving sufficient to maintain the automation condition, output per capita grows at rate $g_D$. The implications of introducing automation at rate $-g_{1-\beta}$ then depend on whether saving is still sufficient to maintain the automation condition. If it is, the per-capita growth rate increases to $g_D + g_{1-\beta}(1-\rho)/\rho$, and the labor share falls to an asymptotic positive value, as observed above. For more on a model of task automation with (something close to) direct labor-augmenting technology growth $g_D > 0$, see Section 4.4.

Now suppose again that $g_D = 0$, but now suppose that saving is not sufficient to maintain the automation condition—as it cannot be if, for instance, *all* tasks become automatable. In this case, some automatable tasks will not be automated. Here, things proceed roughly as in a model of full substitutability. The growth rate is capped at $sA$, the wage rate equals the capital rental rate and stagnates, and the labor share equals $L/(L+K)$. Assuming $sA > g_L$, the labor share falls at rate $g_L - g_K = g_L - sA$.

Finally, consider the implications of exogenous growth in capital-augmenting technology $A$. When not all tasks are automated, increases to $A$ only increase the effective capital stock. If $g_A > 0$ and $s > 0$, even if the automation condition is not yet met, the growth rate $sA$ will grow until it is met. The

[23]The capital share here equals $\beta_t^{1-\rho}(K_t/Y_t)^\rho$. As $\beta \to 1$, the capital share rises to an upper bound of $(K/Y)^\rho$, where $K/Y$ is the long-run capital-to-output ratio, as long as this exists and is finite. It follows from $sY_t = K_{t+1} - K_t$ that $sY_t/K_t = g_{K,t}$; and since $g_K = g_Y = g_B + g_L$, we have $K/Y = s/(g_B + g_L)$. The labor share will thus fall to $1-(s/(g_B+g_L))^\rho$. This is nonnegative because $sA \geq g_B + g_L$, by sufficient saving, and $A$ is asymptotically 1; and it is strictly positive as long as we are not in the knife-edge case of $sA = g_B + g_L$.

[24]One might however take the position that what we have here been calling automation is not a new force on the horizon, promising to augment pre-existing drivers of growth, but a microfoundation for the process of labor-augmenting technological development we have observed for centuries. On this view, advances in AI will continue to push $\beta$ ever closer to 1, but this process will simply continue the existing trend.

capital stock will then grow at the output growth rate, the effective capital stock will grow faster by $g_A$, and the capital share will fall to 0, roughly as explained in Section 3.1.

When all tasks are automated, on the other hand, output $Y_t$ asymptotically equals $A_t K_t$, and sustained exponential growth in $A$ produces a Type I growth explosion.

## 4.3 Task creation

Let us begin with the Aghion et al. (2019) model of automation and introduce a process of task creation. The resulting model will be somewhat akin to that developed by Hémous and Olsen (2014).

As before, output is a CES production of a range of tasks. Each task is performed by labor and/or capital, with tasks above an automation threshold $\beta$ requiring labor. Now, however, new and initially non-automated tasks can be created. The range of tasks thus runs from $i = 0$ to $N$, with tasks $i \leq \beta$ automatable, and not only $\beta$ but also $N$ can be increased. By the same reasoning as in Section 4.1, if there is enough saving that the automation condition is met, output is

$$Y = [(AK)^\rho + (BL)^\rho]^{1/\rho}, \tag{37}$$

where $A = \beta^{(1-\rho)/\rho}$ and $B = (N - \beta)^{(1-\rho)/\rho}$.

If $\rho < 0$, then increases to $\beta$ holding $N$ fixed act like labor-augmenting technology, as in Aghion et al. (2019). By the same token, however, increases to $N$ holding $\beta$ fixed act like labor-depleting technology; they require labor to “spread itself too thinly”. It will never be productive to create new tasks, and automation and growth will simply proceed as in Section 4.2.

If $\rho > 0$, on the other hand, then it is increases to $N$, holding $\beta$ fixed, that function as labor-augmenting technology. In particular, they asymptotically produce $g_B = g_N(1-\rho)/\rho$. As explained in Section 3.2, growth then proceeds at a rate of $\max(sA, g_B)$, where $s$ denotes the saving rate. Increasing the rate of task creation can thus increase the growth rate.

More importantly, increases to $\beta$, regardless of $N$, function as advances in capital-augmenting technology. Recall that effective capital accumulation is enough for growth when $\rho > 0$. By raising the “ceiling” $N$ and allowing for future automation to raise $\beta$, task creation can thus have radical effects.

To see this, first suppose that $N$ increases exogenously at a constant proportional rate $g_N$, and that $N - \beta$ is constant. This is essentially the case explored by Nordhaus (2021) and summarized in Section 3.2: given $\rho > 0$, capital accumulation and capital-augmenting technology growth combine to produce a Type I growth explosion. The labor share falls to 0, even while wages, like output (though more slowly than output), grow superexponentially.

If $g_A = g_B$, i.e. if $g_\beta = g_N$ so that a constant fraction of tasks is always automated, the outcome is similar. Indeed, conditions are even more favorable to labor. Upon endogenizing the task automation and creation processes, the $g_\beta = g_N$ condition turns out to hold under relatively natural-seeming circumstances. For a few more words on this, see the end of the following section.

Finally, as discussed at the end of Section 3.3, suppose we embed production function (37) in a "surrounding" production function. That is, suppose that the good denoted $Y$, which we had been referring to as the final output good, must instead be combined with a fixed-supply resource (such as land) $W$ in order to produce the final output $Z$.

What follows depends centrally, as we have seen, on the substitution parameter $\rho'$ between $Y$ and $W$. If $\rho' > 0$, essentially nothing changes; the land share is asymptotically 0, growth in $Z$ approximately equals growth in $Y$, and so forth. If $\rho' < 0$, on the other hand, output approaches an upper bound. Then the relative quantity of $Y$ rises without bound, by capital accumulation or asymptotic task automation; the land share rises to 1; and the wage rate falls to 0.

This is more similar to the case explored by Hémous and Olsen, though they take the fixed-supply resource to be skilled labor, whereas we are ignoring skill differences throughout this review. In any event, we will not explore this further here.

## 4.4 Task replacement

Acemoglu and Restrepo (2018b, 2019a) develop a similar model of task creation, but combine it with a process of task replacement. Here is a simplification.

Instead of ranging from 0 to $N$, task indices $i$ now range from $N - 1$ to $N$. Capital is equally productive at all tasks it can perform, but labor's productivity at task $i$ is $B_i = D_i(1-\beta)^{(1-\rho)/\rho}$, where $D_i = \exp(g_D i)$, for some $g_D > 0$. In an exogenous growth setting, both $\beta \in (N-1, N)$ and $N$ grow over time at a constant exogenous *absolute* rate—let us say, without loss of generality, at one unit per unit time. The fraction of tasks not automatable is thus constant at $N-\beta$, but the productivity of human labor at the non-automatable tasks grows at exponential rate $g_B = g_D$. With enough saving, all automatable tasks are automated. Output and capital grow at rate $g_D + g_L$, in line with the effective labor supply, and wages grow at rate $g_D$. The labor share is constant, as in any CES model with labor-augmenting technology growth and sufficient saving.

Moving from asymptotic automation in the original setting of Section 4.2 to this model of task automation, creation, and replacement thus increases the output growth rate iff $g_D > g_{1-\beta}(1-\rho)/\rho$. As usual, if we imagine starting from a world without task creation and replacement, introducing this process raises the growth rate from 0 to a positive number; and if we imagine starting from a world with some other source of exogenous growth in labor-augmenting technology, introducing this process raises the growth rate, given enough saving to maintain the automation condition.

This model is nearly equivalent to a task-based model in which the task-range is fixed at the unit interval, $\beta$ is fixed, and $B$ grows exogenously at rate $g_D$. Its framing is motivated by the empirical observation that we have long seen the automation of existing tasks go hand-in-hand with the creation of new, high-productivity tasks that only humans can perform, at least temporarily (Goldin and Katz, 2009; Acemoglu and Autor, 2012). We continue to see this pattern in the recent past (Autor, 2015; Acemoglu and Restrepo, 2019b). The result is a near-complete turnover of job types over time, rather than a mere encroachment of automation onto human territory. In this sense, this promises to be a more realistic model of automation over the past two centuries than one without task replacement. As we have just seen, models of this type are also compatible with balanced growth.

This balanced growth result is, however, sensitive to the assumption that advances in automation technology (increases in $\beta$) and task creation (increases in $N$) proceed at the same rate.

If task creation outstrips automatability, the labor share rises, and as $\beta - N + 1 \to 0$ asymptotically, the labor share rises to 1. In this case, we approach a state in which labor performs all tasks. Capital is relegated to an ever-shrinking band of the lowest-labor-productivity tasks. Since output and, given a constant saving rate, capital grow at the same rate as effective labor, while capital is used ever less efficiently, capital rents fall and the capital share falls to 0. In equilibrium, output grows at rate $g_D + g_L$, and wages grow at the labor-augmenting technology growth rate $g_D$, as before.

Now suppose that automatability outstrips task creation, and in particular that it does so at a constant rate $g_{N-\beta} < 0$. What follows depends on the extent to which capital accumulation keeps up with this process. If $s \leq g_D + g_L$, the automation condition will not be met in the long run, and capital and effective labor will be perfect substitutes on the margin. Output will thus equal $K_t + D_t L_t$, and the stock of capital grows at the same rate as that of effective labor when $s(K_t + D_t L_t)/K_t = g_D + g_L$, i.e. when $K_t/(K_t + D_t L_t) = s/(g_D + g_L)$. In the long run we thus have $g_K = g_Y = g_D + g_L$, the labor share and the fraction of tasks not automated approach $1 - s/(g_D + g_L)$ (ranging from 1 at $s = 0$ to 0 at $s = g_D + g_L$), and wages grow at rate $g_D$.

Now suppose that $s \in (g_D + g_L, g_D + g_{N-\beta}(1-\rho)/\rho + g_L]$, so that capital accumulation outpaces labor and labor-augmenting technology growth in isolation but not in combination with automation. Now the fraction of tasks automated grows over time: if it stayed constant, capital per automated task would ultimately exceed effective labor per non-automated task, since $s > g_D + g_L$, and it would be profitable to reallocate some capital to automatable but non-automated tasks. But the fraction of tasks automated does not catch up with the automatability frontier: if all automatable tasks were automated, effective labor per non-automated task would ultimately exceed capital per automated task, since $s < g_D + g_{N-\beta}(1-\rho)/\rho + g_L$, and it would be profitable to reallocate some labor to currently automated tasks. Thus the automation condition is not met in this scenario either, and capital and effective labor are still perfect substitutes on the margin. Since the stock of capital grows more quickly than that of effective labor, in the long run output grows at rate $s$, wages grow at rate $g_D$, and the labor share again falls to 0.

Finally, if $s > g_D + g_{N-\beta}(1-\rho)/\rho + g_L$, the automation condition is met. Growth proceeds at $g_D + g_{N-\beta}(1-\rho)/\rho + g_L$ and the labor share approaches a positive constant, as in the model of Section 4.2 with direct labor-augmenting technology growth $g_D > 0$. Empirically the saving rate is currently far higher than the growth rate of effective labor, so unless automation accelerates dramatically, this is the most relevant case for consideration.

Now let us briefly and informally consider the implications of endogenizing the automation technology and task creation processes. Suppose that, in addition to the labor force, there is a pool of researchers who allocate their efforts between increasing $\beta$ and increasing $N$. Upon doing either, they earn a patent right to some of the gains that result.

This scenario, as detailed by Acemoglu and Restrepo (2018b), produces intuitive equilibrating pressures, suggesting that we might expect to observe automation technology and task creation proceeding at the same rate, without having to assume this *ad hoc*. Excessive development of automation technology results in tasks that are automatable but not automated, because of an insufficient ratio of capital to effective labor. This eliminates the immediate value of further automation technology. Excessive task creation, on the other hand, increases the value of automation technology, by inefficiently relegating capital to a narrower range of tasks.

The full range of possibilities here, however, is essentially the same as in the exogenous growth case. The proportion of tasks automated can fall to 0, if the saving rate is sufficiently low; there can be asymptotically complete automation if the saving rate is sufficiently high; and there is partial automation in intermediate cases. The primary novelty of the endogenous research case is that here, which case obtains can depend on the researchers' productivity at developing automation technology relative to their productivity at task creation. In particular, increases in productivity at developing automation technology, relative to productivity at task creation, can increase the equilibrium automation rate and decrease the equilibrium labor share. Also, the growth rate here is not exogenous but depends on the level of productivity at both researcher tasks, as well as on the size of the researcher population.

# 5 AI in technology production

Throughout the discussion so far (except for the brief note at the end of the previous section), technological development has been exogenous, when it has appeared at all. Even in this circumstance, developments in AI have proved capable of delivering transformative economic consequences. When technological development is endogenous, and in particular when more advanced AI can allow it to proceed more quickly, the resulting process of "recursive self-improvement" can generate even more transformative consequences.

## 5.1 (Semi-)endogenous growth

Endogenous growth theory models growth in $B$ as the output of a deliberate effort, such as technological research. That is, technology, like final output, is generated from inputs such as labor, capital, and the stock of existing technology. In a standard research-based growth model in the tradition of Romer (1990) and Jones (1995), the growth of $B$ in absolute terms is given by

$$\dot{B}_t = \theta B_t^{\phi}(S_t L_t)^{\lambda} \tag{38}$$

for $\theta > 0$, $\lambda > 0$, and, for our purposes, unrestricted values of $\phi$. $S_t$ denotes the fraction of the labor force working as researchers (or "scientists") at $t$. Intuitively, $\lambda > 1$ corresponds to cases of increasing returns to research, in which researchers complement each other, and $\lambda < 1$ corresponds to cases of decreasing returns, in which some sort of "duplicated work" or "stepping on toes" effect predominates. As before, final output is given by

$$Y_t = F(K_t, B_t(1 - S_t)L_t). \tag{39}$$

Various tweaks to the basic growth model of (38)–(39) are explored throughout the subsections below.

In this setting, though output is CRS with respect to capital and effective labor at any given time, it exhibits *increasing* returns to scale in population across time. We therefore cannot continue to assume that all inputs to production are paid their marginal products. In particular, we cannot assume that technological innovators are compensated for all the additional future output that their research produces on the margin; this sum of marginal products would exceed total output. (Indeed, Nordhaus (2004) estimates that innovative firms accrue on average only about 2% of the value they produce.) It would be beyond the scope of this section to summarize theories regarding the empirical or optimal number of researchers, or the empirical or optimal level of researcher pay. For now, to introduce endogenous technological development without having to consider its interactions with the rest of the framework, we might simply assume that a government sets $S_t$ and pays $S_t L_t$ workers to do research. Their wages are equal to non-research workers' wages, in this stylization, and they are financed by lump-sum taxes levied equally across the population.

To maintain a constant growth rate in output per capita, we must maintain a constant rate of labor-augmenting technology growth, by the reasoning laid out in Section 2.2. However, the growth rate of $B$ at $t$, denoted $g_{B,t}$, is by definition equal to $\dot{B}_t/B_t$. Holding $L$ and $S$ fixed, therefore, we now have

$$g_{B,t} = \theta B_t^{\phi-1}(SL)^{\lambda}. \tag{40}$$

If $\phi < 1$, as we can see, $g_{B,t}$ falls to 0 as $B_t$ grows. If $\phi > 1$, $g_{B,t}$ rises to infinity. Only in the knife-edge case of $\phi = 1$ do we get exponential growth with a constant number of researchers.[25]

Empirically, both the population and the fraction of the population working in research have grown dramatically over the past few centuries. For simplicity, and because the number of researchers cannot grow indefinitely in a fixed population, let us here ignore the second trend and suppose that $S$ is fixed, with $L_t$ growing at a constant rate $g_L$. In this case, labor-augmenting technology growth $g_B$ is constant over time iff

$$B_t^{\phi-1}(SL_t)^{\lambda} = B_0^{\phi-1}e^{g_B(\phi-1)t} \cdot S^{\lambda} \cdot L_0^{\lambda}e^{g_L\lambda t} \tag{41}$$

is constant over time; that is, iff the change in the number of researchers just offsets the change in the difficulty of producing proportional productivity increases. This in turn will obtain iff $(\phi-1)g_B + \lambda g_L = 0$, or[26]

$$g_B = \frac{\lambda}{1-\phi}\, g_L. \tag{42}$$

Because we are holding $1 - S_t$ fixed, the number of non-research workers will grow at rate $g_L$. By extension, the number of effective non-research workers will grow at rate $g_B + g_L$. As we have seen, under a constant saving rate, capital and output will grow at this rate too, and output per person will grow at rate $g_B$.

Observe that the steady-state rate of labor-augmenting technology growth is here undefined when $\phi = 1$. The calculation also breaks down when $\phi > 1$, absurdly predicting a negative rate. This is because the value assumed to exist in the derivation, namely a steady-state technology growth rate $g_B$

[25] Indeed, some reserve the term "endogenous" for growth models in which $\phi = 1$, since an unexplained process of exponential population growth is needed when $\phi < 1$. Models of this form with $\phi < 1$ are then termed "semi-endogenous".

[26] It follows from (40), with $\phi < 1$, that the corresponding growth path is stable. If $B_t$ is "too high", growth subsequently slows, since $\phi - 1 < 0$. Likewise, if $B_t$ is "too low", growth accelerates.

under a growing research workforce, does not exist when $\phi \geq 1$. When $\phi = 1$, it is straightforward to see that a positive growth rate in the number of researchers produces an increasing rate of labor-augmenting technology growth. When $\phi > 1$, even a constant number of researchers produces ever-increasing labor-augmenting technology growth as well.

The ever-increasing $g_{B,t}$ that follows when $\phi = 1$ and $g_L > 0$ translates into an increasing output growth rate up to the point that $g_{B,t} + g_L = sA$. At that point capital accumulation cannot keep up with the growth of the effective labor force, and production is constrained by capital. If capital-augmenting technology can be developed in parallel with labor-augmenting technology, however, the two factors can both grow at an increasing rate, and output therefore can as well. That is, we have a Type I growth explosion.

When $\phi > 1$, moreover, even a constant number of researchers is enough to produce "infinite output in finite time", which Aghion et al. term a Type II growth explosion. The intuition for this is easiest to grasp when $\phi = 2$ and $(SL)^\lambda = 1$, so that we have $g_{B,t} = B_t$. Suppose $g_{B,0}$ is such that technology doubles every time period. Thus $B_1 = 2B_0$, so $g_{B,1} = 2g_{B,0}$. At this doubled growth rate, technology doubles every half-period; $B_{1.5} = 2B_1$. By repeated applications of the same reasoning, the technology level approaches a vertical asymptote at $t = 2$. If capital-augmenting technology follows a similar process, output approaches a vertical asymptote at $t = 2$ as well.

The potential for endogenous growth processes to produce explosive growth is striking. However, since the researcher population growth rate has long been positive and the technology growth rate has long been roughly constant, and in fact declining over recent decades (Gordon, 2016), we can infer that at least historically $\phi < 1$. Indeed, Bloom et al. (2020), the most extensive study of the topic to date, estimates that $\phi \approx -2$. An estimate of $\phi \in (0, 1)$ would indicate that, when we have access to a large stock of existing technologies, these aid in the development of new technologies, but offer diminishing marginal aid. An estimate of $\phi < 0$ implies that when there is a large stock of existing technologies it is *harder* to develop new technologies—perhaps because so much of the low-hanging technological fruit has already been developed, with this "fishing out" effect outweighing the effect of technological assistance in technological development.

## 5.2 Learning by doing

This recursive effect can be seen most simply in a model with Cobb-Douglas production. Let us interpret the production as task-based, with fraction $a$ of tasks automated, as in Section 4.1. The model below is inspired by the exploration of learning by doing in Hanson (2001).

The labor supply grows at exogenous rate $g_L > 0$, but capital-augmenting technology growth proceeds endogenously. We will use a modification of the endogenous growth model presented in Section 5.1. In that model, technology growth is a function of existing technology and "researcher effort". Here, instead, we will not introduce research and will say simply that capital-augmenting technology grows as a function of the existing technology and *output*. That is,

$$Y_t = (A_t K_t)^a L_t^{1-a}, \tag{43}$$

where

$$\dot{A}_t = \theta A_t^{\phi} Y_t^{\lambda} \quad \Rightarrow \quad g_{A,t} = \theta A_t^{\phi-1} Y_t^{\lambda} \tag{44}$$

for some $\theta, \lambda > 0$ and $\phi < 1$. One might interpret this as a model in which the production process itself contributes to the generation of productivity-increasing ideas.

Given a constant saving rate, if growth in the long run is exponential, in the long run we have $g_K = g_Y$. So, from our production function,

$$g_Y = a(g_A + g_Y) + (1-a)g_L \quad \Rightarrow \quad g_Y = \frac{a}{1-a}\, g_A + g_L. \tag{45}$$

From our formula for $g_{A,t}$, the steady state of $g_A$ (if it exists) will be that which satisfies $(\phi - 1)g_A + \lambda g_Y = 0$. Substituting for $g_Y$ in this expression and solving for $g_A$, we have

$$g_A = \frac{\lambda(1-a)}{(1-a)(1-\phi) - \lambda a}\, g_L. \tag{46}$$

This exponential growth path will exist as long as the denominator is positive: that is, as long as

$$a < \frac{1-\phi}{1-\phi+\lambda}. \tag{47}$$

In this case, output growth will be given by[27]

$$g_Y = \frac{(1-a)(1-\phi)}{(1-a)(1-\phi) - \lambda a} g_L. \tag{48}$$

Otherwise, the recursive process by which proportional increases to $A_t$ generate proportional increases to $Y_t$, which in turn generate proportional increases to $A_{t+1}$ (using discrete-time notation for clarity), results in the proportional increases at $t+1$ being larger than those at $t$. The growth rates of $A$ and $Y$ thus increase without bound.

The transformative potential of automation is now straightforward. Increases in $a$ increase the long-run growth rate without bound, as $a$ approaches $(1-\phi)/(1-\phi+\lambda)$. Increasing $a$ past this threshold triggers a growth explosion.

The growth explosion type can be determined by substituting (43) into expression (44) for $g_{A,t}$:

$$g_{A,t} = \theta A_t^{\phi-1} Y_t^{\lambda} = \theta A_t^{\phi-1} A_t^{a\lambda} K_t^{a\lambda} L_t^{(1-a)\lambda}. \tag{49}$$

Since capital is a fraction of all past output, and since here output grows superexponentially, capital grows less quickly than output. An upper bound on $g_{A,t}$, for large $t$, can thus be found by positing that $K_t =$ (43), rearranging to get $K_t = A_t^{a/(1-a)} L_t$, and substituting the result into (49):

$$g_{A,t} < A_t^{\phi-1+\lambda a/(1-a)} L_t^{\lambda}. \tag{50}$$

When $a = (1-\phi)/(1-\phi+\lambda)$, therefore, the technology growth rate itself grows asymptotically at a rate no greater than $\lambda g_L$. The growth explosion is of Type I.

When $a > (1-\phi)/(1-\phi+\lambda)$, a Type II growth explosion obtains even with $L$ fixed. This can be confirmed by observing that the model in this case is the same as that explored in Section 6, after a simple change of variables.

[27]Note that, in this case, effective capital growth $g_A + g_Y$ will always equal $g_L((1-a)(1-\phi)+\lambda(1-a))/((1-a)(1-\phi)-\lambda a) > g_L$. With effective capital growing more quickly than labor, the automation condition will always eventually be met for any fixed $a$.

## 5.3 Automated research

Cockburn et al. (2019) taxonomize AI systems as belonging to three broad categories: symbolic reasoning, robotics, and deep learning. Symbolic reasoning systems, they argue, have proven so far to have few applications. Robotics—by which they broadly mean capital that can substitute for human labor in various ways, instead of complementing it—has had many applications in recent decades. It has also been the subject of a substantial majority of the theoretical literature on the economics of AI, including all that discussed in this survey so far. The most dramatic possibilities, however, may come from deep learning systems that can automate processes of technological development.

Economic reasoning about the transformative effects of automating research goes back at least to Griliches's (1957) discussion of the implications of "inventing a method of invention". The experience of of recent years suggests that deep learning systems are fast becoming methods of invention in the relevant sense. Besiroglu et al. (2023), for instance, present evidence that deep learning has already given rise to significant research automation the field of computer vision. Closer to home, Korinek (2023) evaluates the scope for using generative AI to automate research tasks in economics.

Following Aghion et al. (2019), let us focus directly on the implications of technology production by using an even simpler production function than usual:

$$Y_t = A_t(1 - S)L_t. \tag{51}$$

Labor $L$ is the only factor of production. $S$ is the constant proportion of people who work in research as opposed to final good production. Output technology $A$, however, is developed using a CES function of both labor and capital $K$. Building on the standard technology production function from Section 5.1, i.e., $\dot{A}_t = \theta A_t^{\phi}(S_t L_t)^{\lambda}$, we have

$$\begin{aligned} \dot{A}_t &= A_t^{\phi}[(C_t K_t)^{\rho} + (D_t S L_t)^{\rho}]^{\lambda/\rho}, && \rho \neq 0; \\ \dot{A}_t &= A_t^{\phi}(C_t K_t)^{\lambda a}(D_t S L_t)^{\lambda(1-a)}, && \rho = 0, \end{aligned} \tag{52}$$

for some permitted values of $\rho$, $\lambda$, and, in the Cobb-Douglas case, $a$. (The latter formulation is also the one used by Besiroglu et al. (2023).)

$C_t$ and $D_t$ denote capital- and labor-augmenting technology levels respectively, in the research context. Including them removes the need for a $\theta$

coefficient. The inclusion of factor-augmenting technology terms is unusual in a technology production function, and perhaps somewhat unsatisfying, as it amounts to introducing an explicit technology production function in the final good sector only to leave technology growth in the new "research" sector potentially unexplained. That said, one must introduce a wrinkle along these lines to study the implications of the (exogenous) asymptotic automation of research tasks—keeping in mind the Aghion et al. (2019) result, summarized in Section 4.2, that if $\rho < 0$, asymptotic automation can amount to growth in what is here denoted $D$. The implications of growth in both $C$ and $D$, and across all values of $\rho$, have then been included for generality. To explore the case in which no technology growth is exogenous, simply posit that $C$ and $D$ are constant throughout the discussion below.

As usual, we will assume a constant saving rate, so that capital accumulation in the long run tracks output.

Suppose $\rho < 0$. Recall that in this case sustained growth in effective capital *and* in effective research labor are both necessary to sustain growth in output technology, and growth will be driven by whichever factor grows more slowly. Observe that when growth is constrained by effective capital accumulation, we have[28] $g_A \propto A_t^{\phi-1}(C_t K_t)^\lambda$, and that when it is constrained by effective research labor growth, we have $g_A \propto A_t^{\phi-1}(D_t L_t)^\lambda$.

Regarding $\phi$:

- Recall from the reasoning of Section 5.1 that, if $\phi < 1$, we have $g_A = \lambda(g_C + g_K)/(1-\phi)$ on the capital-constrained path and $g_A = \lambda(g_D + g_L)/(1-\phi)$ on the labor-constrained path. In the former case, since $g_K = g_Y = g_A + g_L$, we can substitute for $g_K$ and rearrange to get $g_A = \frac{\lambda}{1-\phi-\lambda}(g_C + g_L)$. A capital-constrained path with this growth rate exists when $\phi < 1 - \lambda$.

  Thus, if $\phi < 1-\lambda$, $g_A = \min(\frac{\lambda}{1-\phi-\lambda}(g_C + g_L), \frac{\lambda}{1-\phi}(g_D + g_L))$. A one-time increase to $C_t$, $D_t$, $L_t$, or $S$, as long as $S$ remains below 1, does not affect the output technology growth rate. On the other hand, a permanent increase to $g_L$, or to $g_C$ and/or $g_D$, does increase the growth rate of output technology and thereby output per capita.

- If $\phi \in (1-\lambda, 1)$,[29] output technology growth cannot be constrained by

[28] The "$\propto$" symbol means "is asymptotically proportional to".

[29] The dynamics of the knife-edge $\phi = 1 - \lambda$ case are somewhat complex and will be omitted for clarity.

capital accumulation; such a scenario would imply $g_A \propto A_t^{\phi-1}(C_t K_t)^\lambda \propto A_t^{\phi-1+\lambda}(C_t L_t)^\lambda$, contradictorily producing superexponential growth in output technology, output, and capital. We have $g_A = \frac{\lambda}{1-\phi}(g_D + g_L)$.

- If $\phi = 1$, suppose labor remains fixed at $L$ and labor-augmenting technology remains fixed at $D$, while effective capital accumulates. In the long run we then have $g_A = DSL$. A one-time increase to $D$, $S$, or $L$ increases the growth rate of output technology and thereby output (again, as long as $S$ remains below 1). If we begin from a state in which $g_D = g_L = 0$ and introduce positive labor or labor-augmenting technology growth, the result is a Type I growth explosion.
- If $\phi > 1$, we have a Type II growth explosion regardless of the other parameters, as explained in Section 5.1.

As noted above: recall from Section 4.2 that, given $\rho < 0$, increases to $D$ can be interpreted as increases to the fraction of research tasks that have been automated.

Suppose $\rho = 0$. Technology growth is then

$$\begin{aligned} g_{A,t} &= A_t^{\phi-1}(C_t K_t)^{\lambda a}(D_t S L_t)^{\lambda(1-a)} \\ &\propto e^{m_t t}, \end{aligned} \tag{53}$$

where

$$m_t \triangleq g_{A,t}(\phi - 1) + \lambda a(g_C + g_{K,t}) + \lambda(1-a)(g_D + g_L). \tag{54}$$

(Though there is no conceptual distinction between capital- and labor-augmenting technology in the Cobb-Douglas case, both variables have been retained, for easier comparison with the other cases.)

From our assumption of a constant saving rate, exponential growth requires $g_K = g_Y$, which equals $g_A + g_L$. So it requires constant

$$m = g_A(\phi - 1 + \lambda a) + \lambda a g_C + \lambda(1-a)g_D + \lambda g_L. \tag{55}$$

Regarding $\phi$:

- If $\phi < 1 - \lambda a$, we will have, in equilibrium, the constant output technology growth rate that sets $m = 0$. This is $g_A = \frac{\lambda}{1-\lambda a-\phi}(a g_C + (1-a)g_D + g_L)$. One-time increases to $C$, $D$, $S$, or $L$ do not change the growth rate, but increases to $g_C$, $g_D$, or $g_L$ do.

- If $\phi = 1 - \lambda a$, we have steady growth only if $C$, $D$, and $L$ are in the long run constant, since we are assuming that these growth terms are all nonnegative. Fixing $A_0 = K_0 = 1$, the output technology growth rate is $C^{\lambda a}(DSL)^{\lambda(1-a)}$. A one-time increase to $C$, $D$, $S$, or $L$ increases the growth rate. If $C$, $D$, or $L$ grow unboundedly, we have a Type I growth explosion.
- If $\phi > 1 - \lambda a$, we have a Type II growth explosion regardless of the other parameters.

Recall from Section 4.1 that increases to $a$ can be interpreted as increases to the fraction of research tasks that have been automated. They can thus induce Type I and Type II growth explosions, if $\phi \in (0, 1)$.

Suppose $\rho > 0$. Recall that in this case sustained growth in effective capital or in effective research labor suffice to sustain growth in output technology, and growth will be driven by whichever factor grows more quickly. Now, when growth is driven by effective capital accumulation, we have $g_A \propto A_t^{\phi-1}(C_t K_t)^\lambda$, and when it is driven by effective research labor growth, we have $g_A \propto A_t^{\phi-1}(D_t L_t)^\lambda$.

Regarding $\phi$:

- If $\phi < 1 - \lambda$, the capital- and labor-driven technology growth rates equal $\frac{\lambda}{1-\phi}(g_C + g_K)$ and $\frac{\lambda}{1-\phi}(g_D + g_L)$, respectively. In the former case, since $g_K = g_Y = g_A + g_L$, we can substitute for $g_K$ and rearrange to get $g_A = \frac{\lambda}{1-\phi-\lambda}(g_C + g_L)$. Thus $g_A = \max(\frac{\lambda}{1-\phi-\lambda}(g_C + g_L), \frac{\lambda}{1-\phi}(g_D + g_L))$. Growth rate increases require increases to $g_C$, $g_D$, or $g_L$.
- If $\phi > 1 - \lambda$,[30] we have a Type II growth explosion regardless of the other parameters.

An intuition for these results is as follows. As explained briefly in Section 5.1, the growth of some variable $X$ exhibits a Type II growth explosion if its growth rate takes the form $g_X \propto X^\psi$ for some $\psi > 0$. When $\rho < 0$, capital accumulation cannot accelerate technological development, which is bottlenecked by its slower-growing factor, namely effective labor. Output technology growth is then $g_A \propto A^{\phi-1}$, so the Type II growth explosion requires $\phi > 1$. When $\rho > 0$, on the other hand, capital accumulation in line

[30] As under $\rho < 0$, the $\phi = 1 - \lambda$ case has been omitted for simplicity.

with (albeit lagged) output and technology effectively multiplies $g_A$ by a factor of $A^\lambda$, so $g_A \propto A^{1-\phi+\lambda}$. The Type II growth explosion therefore requires only $\phi > 1 - \lambda$.

Note that our analysis of the $\rho > 0$ case also covers the case in which the development of output technology is fully automated. Simply use $\rho = 1$.

It also covers a common interpretation of the possibility of "recursive self-improvement". If $A$ represents cognitive ability, and enhanced intelligence (human or artificial) speeds the rate at which intelligence can be improved such that $g_{A,t} \propto A_t^{\phi-1} K_t^\lambda$, then explosive growth obtains iff $\phi > 1 - \lambda$ (since, again, $K$ grows in line with $A$, in a relevant sense). If we remove capital accumulation entirely and say that the development of intelligence depends only on the intelligence level, such that $g_{A,t} \propto A_t^{\phi-1}$, then explosive growth obtains iff $\phi > 1$.

## 5.4 AI assistance in research

In Section 4, we discussed several papers which use a microfoundation of the output production function as a basis for exploring the implications of a certain kind of automation. Somewhat analogously, Agrawal et al. (2019) use Weitzman's (1998) "combinatorial" microfoundation of the process of technological development as a basis for exploring the implications of a certain way in which advances in AI might assist in technological development.

Let $Y = A(1-S)L$, as before, and hold $S$ fixed but posit labor growth $g_L > 0$. Given $A$ existing "technological ideas", a researcher has access to only $A^\phi$, for some $\phi \in (0, 1)$, perhaps due to some sort of cognitive limitation.[31] New ideas are made from combinations of existing ideas. Given access to $A^\phi$ ideas, a researcher therefore faces $2^{A^\phi}$ idea-combinations. Of these, not all can generate new technological ideas, perhaps due to some other sort of cognitive limitation. Instead, each researcher's idea-generation function is "isoelastic" in ideas available:

$$\begin{aligned} \dot{A} &= \theta \frac{(2^{A^\phi})^\alpha - 1}{\alpha}, \ \alpha > 0; \\ \dot{A} &= \theta \ln(2^{A^\phi}) = \theta \ln(2) A^\phi, \alpha = 0, \end{aligned} \tag{56}$$

[31]The model requires $\phi > 0$ such that the fishing-out effect does not predominate. As discussed in Section 5.1, Bloom et al. (2020) estimate $\phi \approx -2$.

for some $\theta > 0$ and some $\alpha \in [0, 1]$.[32]

Suppose that $\theta = 0$ (or that $\theta \to 0$ as $A \to \infty$), and that total research output is linear in the number of researchers raised to some positive power $\lambda$, as in the growth model of Section 5.1. Then collective technological development is given (at least asymptotically) by

$$\dot{A} = \theta \ln(2) A^{\phi} (SL)^{\lambda} \;\; \Rightarrow \;\; g_A = \frac{\lambda g_L}{1 - \phi}. \tag{57}$$

This is just the standard Jones model, with a coefficient of ln(2) rescaling $\theta$.

Now suppose that $\alpha > 0$ (or that $\alpha$ is bounded below by $\underline{\alpha} > 0$ as $A \to \infty$). Then collective technological development is bounded below (at least asymptotically) by

$$\begin{aligned} \dot{A} &= \theta \frac{(2^{A^{\phi}})^{\underline{\alpha}} - 1}{\underline{\alpha}} (SL)^{\lambda} \\ \Rightarrow g_A &= \theta \frac{(2^{A^{\phi}})^{\underline{\alpha}} - 1}{A\underline{\alpha}} (SL)^{\lambda}. \end{aligned} \tag{58}$$

It follows from the second term that, for large $A$, $g_A$ increases more than polynomially in $A$. That is, $g_A$ increases quickly enough in $A$ to produce a Type II growth explosion. If the above model approximates reality, therefore, we presumably currently have $\alpha = 0$.

Given $\alpha = 0$, let us now consider the potential impacts of artificial research assistance.

If it allows for a one-time increase to $\theta$, this amounts to a one-time increase to the supply of effective researchers. This puts us on a higher growth path, but it does not increase the growth rate or have any other transformative effects. But if AI assistance improves with time in this way, allowing for $\theta_t$ to grow at some positive exponential rate, this amounts to an increase in the growth rate of effective researchers. It can thus increase the growth rate of technology and thereby output.

If AI assistance allows researchers to access more of the stock of existing ideas, it amounts to a one-time increase in $\phi$. (One might argue that this is what internet library access and accurate search engines have already enabled.) As we can see, this increases the growth rate as well.

[32] The formula for $\alpha = 0$ is the limiting case of the formula for $\alpha > 0$, as $\alpha \to 0$.

Most transformatively, if AI tools help researchers search through the ever-growing "haystacks" of possible idea-combinations for valuable "needles", they could permanently increase $\alpha$. Agrawal et al. (2019) argue that this is precisely the sort of activity to which AI systems are best suited: they are already being profitably used to identify promising combinations of chemicals in pharmaceutical development, for example. (See Agrawal et al. (2018) for a more thorough defense of this argument.) If this turns out to hold across the board, the result is stark: as shown above, a permanent increase to $\alpha$ produces a Type II growth explosion.

Agrawal et al. (2019) also explore the potential impacts of AI assistance in research teams, rather than in assisting individual researchers. Seeber et al. (2020) do the same, in the context of a much more applied and less formal inquiry. Neither analysis appears to reveal channels for transformative growth effects substantively different from those presented above.

## 5.5 Growth impacts via impacts on technology investment

Throughout Section 5, we have taken technology production to be endogenous, in the sense that it has required explicit inputs of capital and labor. Nevertheless, we have taken the *drivers* of investment in technological development—the fraction $S$ of labor, and (in Section 5.3) the amount of capital, allocated to research—to be exogenous. A final way in which AI could have a transformative impact, therefore, is by changing the levels of investment in, and effort allocated to, technological development. As we have seen, at least in some circumstances, this can change the growth rate, or can determine the existence or type of a growth explosion.

This pathway to transformative impact is somewhat analogous to the possibility, explored in Section 3.4, that developments in AI could affect the growth rate by affecting the saving rate, even in an economy without endogenous technological development. As in that case, this change could in principle be positive or negative. (One way AI could impact the extent to which resources are devoted to technological development *is* by affecting the saving rate, as long as capital is modeled as an input to technology production.) Also as in that case, to the extent that the literature has explored this pathway to transformative impact, it has focused on the perhaps counterintuitive possibility that AI slows growth.

This could take place by accelerating the “Schumpeterian” process of “creative destruction”. In this analysis, the incentive to innovate comes from a temporary monopoly that the innovators enjoy, either by patents or by trade secrets, during which they can extract rents from those who would benefit by using the new technology in production. AI, however, could make it easier for competitors to copy innovations. Relatedly, AI could also ease the rapid development of technologies only negligibly more productivity-enhancing than those they replace. Because these technologies would entirely eliminate the markets for those they replace, their rapid development would curtail the incentive for innovation. In the absence of this incentive, technology growth can slow to a halt. This cannot cause output per capita to fall, at least in most models, but it can cause output per capita to stagnate.

This dynamic is explored more formally by Aghion et al. (2019) in the context of the model of automated research laid out in Section 5.3, and by Acemoglu and Restrepo (2018b) in the context of the model of automation and task replacement laid out in Section 4.4. We will not work through it here.

As with the Sachs and Kotlikoff (2012) observation that AI can do damage by lowering the saving rate, the insight here is not primarily an insight about artificial intelligence. It is primarily a special case of the well-known fact, mentioned briefly in Section 5.1, that though free and competitive markets can generally be expected to appropriately compensate production factors for a final good in a static setting, the same cannot be said about the inputs to technological development. Policymakers interested in growth must always consider the impact of structural economic changes on the incentives for technological innovation, therefore, and must adjust their funding or subsidization of basic research in light of such changes as they unfold.

# 6 AI in both good and technology production

Naturally, the effects of AI are most transformative of all when it allows capital to better substitute for labor in both good production *and* technology production. Unfortunately, this pair of circumstances has been studied even less extensively than the effects of AI in each sector separately. Nevertheless, an analysis that begins with the research automation of Section 5.3, but replaces the labor-only production function with a CES one, proves relatively straightforward.

Suppose we replace the labor-only final good production function, (51) from Section 5.3, with a CES production function in capital and labor. Let the substitution parameter in the final good sector be denoted $\rho_Y$, and that in the research sector be denoted $\rho_A$. We will ignore factor-augmenting technology in the good production function; output technology $A$ will be thought of as augmenting both. We will assume a constant and sufficient saving rate $s$, and fractions of capital and labor used in the technology sector—$S_K$ and $S_L$ respectively—strictly between zero and one.

We can now consider the growth regimes that obtain under different values of $\rho_Y$ and $\rho_A$. For simplicity, we will not allow for labor growth or exogenous sources of technology growth. To begin, let us list the cases we have already implicitly covered.

If $\rho_Y < 0$, little changes from the case of Section 5.3. Output is still bottlenecked by the scarce factor, namely labor. Output therefore asymptotically resembles $A(1 - S_L)L$, as before. How technology evolves depends on $\rho_A$, as covered in Section 5.3.

If $\rho_Y = 0$ but $\rho_A < 0$, we are in the well-worn territory of Cobb-Douglas production—so, given capital depreciation, production per capita that grows with technology—and technology that grows sub-exponentially (unless research labor inputs grow exponentially).

If $\rho_Y > 0$ but $\rho_A < 0$, we reach the Type I growth explosion discussed in Section 3.2. Output in the absence of growth in $A$ grows at rate $sA$, but $A$ grows without bound, even without growth in research labor inputs.

If $\rho_Y = \rho_A = 0$, however, the ability of capital to contribute to both good production and technology production generates possibilities we have not yet considered. As we will see, the resulting growth path is highly sensitive to the other parameters.[33]

In particular, let

$$Y_t = A_t((1 - S_K)K_t)^a, \tag{59}$$

$$\dot{K}_t = sY_t, \text{ and} \tag{60}$$

$$\dot{A}_t = \theta A_t^{\phi}(S_K K_t)^{\lambda}, \tag{61}$$

[33]What follows is an elaboration on Aghion et al. (2019), Section 4.1, Example 3.

where $S_K \in (0,1)$, $a \in (0,1)$, $s > 0$, $\theta > 0$, $\phi < 1$, and $\lambda > 0$. Also, define

$$\gamma \triangleq \frac{\lambda}{(1-a)(1-\phi)}. \tag{62}$$

Then

- If $\gamma > 1$, $Y$ exhibits a Type II growth explosion.
- If $\gamma = 1$, $Y$ grows exponentially, with

$$\lim_{t\to\infty} g_{Y,t} = s^{\frac{\lambda}{1+\lambda-a}} \left(\frac{\theta}{1-a} \frac{S_K^\lambda}{(1-S_K)^a}\right)^{\frac{1-a}{1+\lambda-a}}. \tag{63}$$

- If $\gamma < 1$, $Y$ grows power-functionally.

Note that the production function of (59) is Cobb-Douglas with an implicit constant labor stock normalized to 1, and/or a constant land stock also normalized to 1, and (given CRS) the exponents on labor and land summing to $1-a$. Note likewise that technology production, as described by (61), may be interpreted as Cobb-Douglas with inputs other than capital fixed.

A proof of the above can be found in Appendix B.2, but an intuition for the exponential growth threshold provided by $\gamma = 1$ is as follows. If growth in $A$ and $Y$ were driven by exogenous exponential growth in $K$, we would have, in steady state,

$$g_A = \frac{\lambda}{1-\phi} g_K \tag{64}$$

and thus

$$g_Y = \left(\frac{\lambda}{1-\phi} + a\right) g_K. \tag{65}$$

But $g_K$ is not exogenous: future growth in $K$ roughly equals past growth in $Y$, since capital accumulation is driven by saving a proportion of output. If $\frac{\lambda}{1-\phi} + a > 1$, therefore, a given growth rate in $K$ generates a higher growth rate of $Y$, and this higher growth rate is subsequently exhibited by $K$. The growth rate of $K$ therefore grows over time. Likewise, if $\frac{\lambda}{1-\phi} + a < 1$, a given growth rate in $K$ generates a lower growth rate of $Y$, and this lower growth rate is subsequently exhibited by $K$. The growth rate of $K$ therefore falls. Finally, observe that $\frac{\lambda}{1-\phi} + a > 1$ iff $\gamma > 1$, and likewise for $<$ and $=$.

If we replace the $\rho_Y = \rho_A = 0$ model with one in which (a) $\rho_Y > 0$ and/or (b) $\rho_A > 0$, nothing changes except that, respectively,

1. the exponent on capital in good production effectively rises from $a$ to what, in a fully specified Cobb-Douglas model, would have been the sum of the exponents on capital and labor in good production; and/or
2. the exponent on capital in idea production effectively rises from $\lambda$ to what, in a fully specified Cobb-Douglas model, would have been the sum of the exponents on capital and labor in idea production.

Let us denote these new exponents $\tilde{a}$ and $\tilde{\lambda}$.

In the absence of a natural resource constraint, $\tilde{a} = 1$, by the assumption of CRS good production. Since, from (62), we have $\lim_{a \uparrow 1} \gamma = \infty$, it follows that, absent significant natural resource constraints, $\rho_Y > 0$ and $\rho_A \geq 0$ always produce a Type II growth explosion.

By contrast, we do not in general assume that $\tilde{\lambda} = 1$, i.e. that research outputs exhibit constant returns to scale in research inputs. Furthermore, even if we did, a value of $\lambda = 1$ is not sufficient (or, for that matter, necessary) for $\gamma > 1$. An assumption of $\rho_A > 0$ therefore has no qualitative implications beyond those of the $\rho_A = 0$ case.

# 7 Overview of the possibilities

The table below summarizes the transformative scenarios we have considered. They have been rearranged slightly for clarity, and some near-redundant possibilities have been removed, but they primarily follow the order in which they are presented in Sections 3 to 6. Relevant literature is cited below each scenario. Note that, in keeping with the presentation so far, the cited literature introduces the models that allow for the scenarios in question, but does not always discuss the transformative scenarios on which we have focused.

We have not considered all possible AI scenarios, as this table makes clear. Nevertheless we have hopefully sampled the possibilities thoroughly enough that the reader is now comfortable filling some of the gaps.

| **Scenario**[34] | **Growth**[35] | **Labor share**[36] | **Wages**[37] |
|---|---|---|---|
| LS in production & capital-augmenting tech growth<br>Section 3.1 Acemoglu and Restrepo (2018a) | $=$ | $\rightarrow \mathbf{1}$ | $+$ |
| HS in consumption goods<br>Section 3.2 Nordhaus (2021) | $++$ | **C** | $++$ |
| HS in production<br>Section 3.2 Nordhaus (2021) | $++$ | $\rightarrow \mathbf{0}$ | $++$ |
| HS (not PS) in production & capital-augmenting tech growth<br>Section 3.2 Nordhaus (2021) | **I** | $\rightarrow \mathbf{0}$ | **I** |
| PS in production & capital-augmenting tech growth<br>Section 3.2 | **I** | $\rightarrow \mathbf{0}$ | **L** |
| PS in production, capital-augmenting tech growth, & MS land constraint<br>Section 3.3 Hanson (2001) | $++$ | $\rightarrow \mathbf{0}$ | $\rightarrow \mathbf{0}$ |
| PS in production, equipment-augmen- ting tech gr., & MS land constraint<br>Section 3.3 | $++$ | $\rightarrow \mathbf{0}$ | **L** |

[34] "PS", "HS", "MS", and "LS" stand for perfect, high, moderate, and low substitutability, and refer to substitution parameters $\rho = 1$, $> 0$, $= 0$, and $< 0$ respectively. Unless otherwise noted, the "HS" case allows for perfect substitutability. In the scenarios with endogenous research, "negative", "positive", "low", "intermediate", and "high [research] feedback" refer to research feedback exponents $\phi < 1 - \lambda$, $> 1 - \lambda$, $< 1$, $= 1$, and $> 1$ respectively.

[35] $+$ and $-$ refer to cases in which AI shifts the output path up or down without changing the growth rate, e.g. by increasing or decreasing the plateau level in a circumstance where output plateaus regardless of AI. **--**, $++$, **I**, and **II** refer to cases in which AI allows for decreases to the long-run growth rate, increases to the long-run growth rate, Type I growth explosions, and Type II growth explosions. $=$ refers to cases in which AI does not change the long-run output level or growth rate.

[36] **C** means that AI pushes the human labor share to some positive constant, not necessarily lower or higher than the value it would take in the absence of AI.

[37] **L** means that human wages are driven to some low but constant rate (typically the rental rate of effective capital). **C** means that they are pushed to some positive constant, not necessarily lower or higher than they would be in the absence of AI. All other symbols are defined as in the Growth column.

| | | | |
|---|---|---|---|
| PS in production & LS land constraint (regardless of tech)<br>Section 3.3 Korinek and Stiglitz (2019) | = | $\to$ **0** | $\to$ **0** |
| HS in final good production, HS in robotics production<br>Section 3.4 Mookherjee and Ray (2017), Korinek and Stiglitz (2019) | ++ | $\to$ **0** | **L** |
| HS in final good production, LS in robotics production<br>Section 3.4 Mookherjee and Ray (2017), Korinek (2018) | + | **C** | **C** |
| PS in production & one-off capital-augmenting tech increase $\to$ saving increase<br>Section 3.5 Korinek and Stiglitz (2019) | ++ | $\to$ **0** | = |
| PS in production & capital-aug. tech growth $\to$ saving decrease<br>Section 3.5 Sachs and Kotlikoff (2012), Sachs et al. (2015) | − | $\to$ **0** | $\to$ **0** |
| MS in production & asymptotic or full task automation<br>Section 4.2 Aghion et al. (2019) | **I** | $\to$ **0** | **I** |
| LS in production & asymptotic task automation<br>Section 4.2 Aghion et al. (2019) | ++ | **C** | ++ |
| LS in production & task automation and replacement<br>Section 4.4 Acemoglu and Restrepo (2018b) | ++ | **C** | ++ |
| HS in production & task automation and creation<br>Section 4.3 Hémous and Olsen (2014) | **I** | $\to$ **0** | **I** |
| Learning by doing, w/intermed. feedback and/or automation<br>Section 5.2 Hanson (2001) | ++ | | |
| Learning by doing, with suffic. feedback and/or automation<br>Section 5.2 Hanson (2001) | **II** | | |
| LS in tech production, low research feedback, & asymptotic research task automation; or HS in tech production, negative research feedback, & research capital tech growth<br>Section 5.3 Aghion et al. (2019) | ++ | | |

| | |
|---|---|
| LS in tech production, intermed. research feedback & asymp. research task automation; or HS in tech prod., zero research feedback, & research capital tech growth<br>Section 5.3 Aghion et al. (2019) | **I** |
| LS in tech production & high research feedback or HS in tech production & positive research feedback<br>Section 5.3 Aghion et al. (2019) | **II** |
| AI-assisted multiplication of combinatorial idea discovery<br>Section 5.4 Agrawal et al. (2019) | ++ |
| AI-assisted elasticity-change in idea discovery<br>Section 5.4 Agrawal et al. (2019) | **II** |
| AI-diminished innovation incentives<br>Section 5.5 Aghion et al. (2019), Acemoglu and Restrepo (2018b) | −− |
| HS in production & MS or HS in idea production (for any value of research feedback)<br>Section 6 Aghion et al. (2019) | **II** |

The human labor share and wage are technically undefined in the models of endogenous technology production, since—as noted in Section 5.1—we cannot straightforwardly assume that the factors of technology production will tend to be paid their marginal products (or anything else in particular). As often presented, however, human labor is the lone factor of final good production in these models, and the technology being produced is labor-augmenting. We can therefore assume that the wage rate corresponds to the marginal product of labor in final goods production and grows in line with technology and so with output. That is, it should exhibit growth rate decreases, increases, Type I growth explosions or Type II growth explosions as listed above.

# 8 Conclusion

The literature we synthesized covers a wide range of AI's possible long-run macroeconomic impacts. It can hopefully serve as a bridge between the tools of economics and the larger-scale questions posed by futurists.

Our most significant conclusion is that significant changes to the growth regime experienced by our economies are possible across a range of models, including permanent shifts in growth rates and singularities that involve ever-increasing growth rates. As we illustrate, the lack of attention to these possibilities by most of the economics profession is not a necessary consequence of all plausible economic models. Rather, it is the result of a widespread norm of focusing only on model scenarios in which long-run growth is roughly constant—in line with the Kaldor Facts that described growth over the past century or two. Even Aghion et al. (2019), who take the growth potential of AI most seriously, focus less on scenarios in which labor and capital are highly substitutable in technology production *on the grounds* that "researchers are not a necessary input and so standard capital accumulation is enough to generate explosive growth. This is one reason why the case of [gross substitutes] is the natural case to consider." Expressed motivations along these lines appear throughout the literature.

In summary, there is no shortage of mechanisms through which advances in automation could have transformative growth consequences, once we allow ourselves to look for them. Taking the Industrial Revolution as an example, there is also historical precedent for significant changes in growth regimes.

There are many topics relevant to the economics of AI and transformative AI that we could not cover here.

Wage distribution is a central concern of the literature on the long-term economic implications of AI, including much of the literature cited here. It is likewise a central concern of the less long-term-focused reviews of the economics of AI cited in Section 1. Wages and workers' skill levels are highly unequal, and this inequality has increased in recent decades—a development that many attribute to the rise of automation (e.g. Acemoglu and Restrepo, 2022).

We focus on average wages and on the overall labor share in part for focus, but also in part because in the long run, the likeliest transformative scenario is that AI will outsmart us all. In this event, individual human talents will not save us; if we retain positive wages or a positive labor share, we will do so only because AI is put to use making us more productive, or because we actively decide to keep some jobs, like clergy or hospice nurses, unautomated. To be clear, however, this is not the only scenario. If labor remains important as a means of income distribution in the long run, AI may increase income concentration or—in more optimistic scenarios—could be used to generate greater shared prosperity (Klinova and Korinek, 2021).

Another important subject that is absent from our discussion but that features heavily in futurist discussion about AI is the most transformative macroeconomic possibility of all: the risk of an AI-induced existential catastrophe (see e.g. Bostrom, 2014). With the recent exception of Jones (2023), treatments of existential risk from AI appear to be absent from the economics literature.[38] This is not primarily "by choice" but because there is no particularly obvious mechanism through which accelerating automation, within existing models of production or growth, can pose a danger. In the literature on AI safety, such concerns typically arise from superintelligent agents, with goals not fully aligned with ours, who take control of the world. Even though economic growth in such a scenario could continue without humans, such an outcome is certainly not in our interest.

The tools of economics can help to shed light on these concerns as well. Most simply, to the extent that AI development poses such a risk, AI safety is a global and intergenerational public good. Through that lens, much of the analysis of public goods, and in particular many of the tools developed by environmental economists for the pricing and provision of climate risk mitigation, could apply to AI safety.

More subtly, to the extent that AI risk arises from the ability of AI systems to control resources independently of human input, models in which the labor share remains positive and significant should give us comfort. If human work remains a bottleneck to growth—say, if AI accelerates growth but continues to rely on human workers to perform certain physical tasks—then humans remain essential. More worrying are models in which a unit of capital can grow, do research into capital-augmenting technology, and recursively self-improve, all without human input.

A thorough analysis of the links between the economics of AI and AI safety remains an important topic for further research.

[38]Risk from AI is sometimes listed among examples of catastrophic risks to motivate more generic models of catastrophic risk, as in Martin and Pindyck (2015), Aschenbrenner (2020), or Acemoglu and Lensman (2023).

# A Elasticity of substitution

Given a production function and a list of factor prices, suppose an economic agent spends a fixed budget so as to maximize output. Suppose furthermore that the production plans are CRS. The elasticity of substitution between two factors $i$ and $j$ is then, intuitively, the value $\epsilon$ such that, if the relative price of $i$ falls by a small proportion (say, 1%), the relative quantity of $i$ purchased $X_i/X_j$ rises by an $\epsilon$-times larger proportion (say, $\epsilon$%).

$$\epsilon = -\frac{\partial \log(X_i/X_j)}{\partial \log(p_i/p_j)} = -\frac{\frac{\partial(X_i/X_j)}{X_i/X_j}}{\frac{\partial(p_i/p_j)}{p_i/p_j}}$$

The second equality holds since we can express the derivative of a logarithmic function in terms of the derivative of its argument, i.e., $\frac{d \log X_i}{da} = \frac{dX_i/da}{X_i}$, for any positive and differentiable function $X_i(\cdot)$.

The elasticity of substitution is defined analogously for utility: given a utility function that is homothetic (i.e., that is CRS, or a monotonic transformation of one that is CRS) and a list of prices of consumption goods, it captures how much a relative change in prices affects the relative quantity of goods consumed.

Conversely, given a list of factor (or consumption good) *quantities*, the elasticity of substitution is the value $\epsilon$ such that, in order for the relative quantity of $X_1$ sold to increase by a small proportion (say 1%), the relative price of $X_1$ must fall by a $(1/\epsilon)$-times larger proportion ($(1/\epsilon)$%). For example, suppose that goods $X_1$ and $X_2$ are traded in a competitive market so that their prices equal the market-clearing prices. Now consider the consequences of exogenously increasing the relative supply of $X_1$ by 1%, say by increasing the endowment of each factor owner before trade takes place. The market-clearing relative price of $X_1$ will then fall by $(1/\epsilon)$%. As we can see, marginally increasing the relative abundance of a good results in smaller relative expenditure on that good—i.e. its owners receive a smaller share of total income—precisely when the elasticity of substitution between it and other goods, on the current margin, is less than 1 because $1/\epsilon > 1$ in that case.

For illustration, food and other goods are typically not very substitutable. When food was much *scarcer*, its owners were able to command such higher prices for it that people spent larger shares of their incomes on it. On the other hand, industrially produced goods and handmade goods are very substitutable. As the former grew *more plentiful*, following the Industrial Revolution's explosion in manufacturing, people spent larger shares of their incomes on them.

Goods are perfect complements if $\epsilon = 0$. In this case, output (or utility) is some constant times the minimum of the goods' quantities, in some ratio. (Consider left shoes and right shoes, in the 1:1 ratio, or bicycle frames and wheels, in the 1:2 ratio.) Whatever the way in which the relative prices of such goods change, the relative quantities purchased will stay fixed at the given ratio.

The case of perfect substitutability is approached in the limit as $\epsilon \to \infty$. In this case, a positive quantity of each good is only purchased if their prices are equal; if their prices differ, only the cheaper one is purchased.

For ease of notation, define the "substitution parameter" $\rho \triangleq (\epsilon - 1)/\epsilon$. Note that the cases $\epsilon < 1$, $\epsilon = 1$, and $\epsilon > 1$ correspond to $\rho < 0$, $\rho = 0$, and $\rho > 0$ respectively.

A production function exhibits constant elasticity of substitution (CES) if its elasticity of substitution is constant and does not depend on the factor prices and quantities. It can be shown that any two-factor CES production function that is also CRS and allows for factor-augmenting technology must take the form

$$Y = [(AK)^\rho + (BL)^\rho]^{1/\rho} \tag{66}$$

if $\rho \neq 0$ and

$$Y = (AK)^a (BL)^{1-a}, \tag{67}$$

for some $a \in (0, 1)$, if $\rho = 0$. In the second case, the function is called "Cobb-Douglas".

A common alternative for the parameter $\rho$ is the elasticity of substitution of a production function, $\sigma = 1/(1 - \rho)$, which captures by what percentage the ratio of the two input factors capital and labor changes in response to a one percent change in the marginal rate of technical substitution (MRTS), which, in equilibrium, equals the relative price of the two factors.

More formally, the elasticity of substitution ($\sigma$) is given by:

$$\sigma = -\frac{d \ln(L/K)}{d \ln(\text{MRTS})}$$

where MRTS $= -\frac{F_L}{F_K}$ reflects the ratio of marginal products of labor and capital. In a competitive economy, this ratio equals the relative factor prices of labor and capital.

When $\sigma < 1$ ($\rho < 0$), we call the two factors gross complements. For $\rho \leq 0$, output requires strictly positive quantities of both factors. When $\sigma > 1$ ($\rho > 0$), we call the two factors gross substitutes. Output can be produced even if the input of one of the two factors is zero, i.e., the two factors are no longer *essential*.

When $\rho \neq 0$, the share of output paid to factor $X$, with factor-augmenting technology $C$, equals $(CX/Y)^\rho$. When $\rho = 0$, a factor's share equals the exponent on that factor, i.e., $a$ and $1 - a$. In general, the share of factor $X$ is decreasing in $CX/Y$ when $\rho < 0$, independent of $CX/Y$ when $\rho = 0$, and increasing in $CX/Y$ when $\rho > 0$.

# B Proofs

## B.1 Asymptotically positive fractions of capital and robotics used in robotics production

We will work within the framework of Section 3.4.

Consider a time $t$ at which $S_{R,t} > 0$, and let $m > 1$ denote $K_{t'}/K_t$ for some $t' > t$. From $t$ to $t'$, the capital input to robotics production is multiplied by $mS_{K,t'}/S_{K,t}$. Because $f(\cdot)$ is CRS, to maintain the condition that $f_{L,t'} = f_{L,t} = 1/D$ the labor input to robotics production must also be multiplied by $mS_{K,t'}/S_{K,t}$, and robotics production will then also be multiplied by this factor. We thus have

$$H + DS_{R,t'}R_{t'} = (H + DS_{R,t}R_t)mS_{K,t'}/S_{K,t} \quad \text{and} \tag{68}$$

$$R_{t'} = R_t mS_{K,t'}/S_{K,t}. \tag{69}$$

Because both inputs to robotics production are multiplied by a common quantity, $f_K$ is constant across periods. It follows that if the capital input to final good production grows proportionally more (less) than the labor input, the marginal productivity of capital in final good production falls (rises), and the marginal contribution of capital to final good production via robot production rises (falls). Thus, to maintain the condition that capital is allocated efficiently, the capital and labor inputs to final good production must be multiplied by a common quantity across periods:

$$m\,\frac{1-S_{K,t'}}{1-S_{K,t}}=\frac{R_{t'}}{R_t}\,\frac{1-S_{R,t'}}{1-S_{R,t}}. \tag{70}$$

Substituting (69) into (68) and (70) and solving for $S_{K,t'}$ and $S_{R,t'}$, we find that, as $m\to\infty$,

$$S_{K,t'}\to S_K^*\triangleq S_{K,t}\,\frac{DR_t(1-S_{R,t})}{DR_t(1-S_{R,t})-(1-S_{K,t})H}\ \text{and} \tag{71}$$

$$S_{R,t'}\to S_R^*\triangleq S_{R,t}+H/(DR_t). \tag{72}$$

$S_K$ and $S_R$ are thus asymptotically constant and nonzero. Furthermore, since $R_t=f_{K,t}S_{K,t}K_t+(H+S_{R,t}DR_t)/D$, we must have $H+S_{R,t}DR_t<DR_t$. It follows that $S_K^*$ and $S_R^*$ are strictly less than 1.

## B.2 Growth paths given Cobb-Douglas production and research

As in Section 6, suppose

$$Y_t=A_t((1-S_K)K_t)^a, \tag{73}$$

$$\dot K_t=sY_t,\ \text{and} \tag{74}$$

$$\dot A_t=\theta A_t^{\phi}(S_K K_t)^{\lambda}, \tag{75}$$

where $A_0>0$, $K_0>0$, $S_K\in(0,1)$, $a\in(0,1)$, $s>0$, $\theta>0$, $\phi<1$, and $\lambda>0$, and where (73)–(75) are defined for $t\in[0,\infty)$—or, if the system exhibits a Type II growth explosion at some time $t^*$, for $t\in[0,t^*)$.

Observe first that, for all $t$,

$$g_{Kt}=s(1-S_K)^aA_tK_t^{a-1}\ \text{and} \tag{76}$$

$$g_{At}=\theta S_K^{\lambda}A_t^{\phi-1}K_t^{\lambda}. \tag{77}$$

Let $\hat g_K$ ($\triangleq$"$g_{g_K}$") denote the proportional growth rate of $g_K$ itself, and let $\hat g_A$ be defined likewise. It then follows from (76) and (77) that, for all $t$,

$$\hat g_{Kt}=g_{At}+(a-1)g_{Kt}\ \text{and} \tag{78}$$

$$\hat g_{At}=(\phi-1)g_{At}+\lambda g_{Kt}. \tag{79}$$

If, for <u>any</u> time $\tau$, $\hat{g}_{K\tau} > 0$ and $\hat{g}_{A\tau} > 0$, then

$$\begin{aligned} g_{A\tau} + (a-1)g_{K\tau} &> 0 \\ \implies g_{A\tau} &> (1-a)g_{K\tau}\ ; \qquad (80) \\ (\phi - 1)g_{A\tau} + \lambda g_{K\tau} &> 0 \\ \implies g_{K\tau} &> \frac{1-\phi}{\lambda} g_{A\tau}\ ; \qquad (81) \end{aligned}$$

and thus

$$\begin{aligned} g_{A\tau} &> \frac{(1-a)(1-\phi)}{\lambda} g_{A\tau} \qquad (82) \\ \implies \gamma &> 1 \qquad (83) \end{aligned}$$

since $g_{A\tau} > 0\ \forall \tau$ by construction.

Likewise, if for any $\tau$ we have $\hat{g}_{K\tau} < 0$ $(= 0)$ and $\hat{g}_{A\tau} < 0$ $(= 0)$, then $\gamma < 1$ $(= 1$, respectively).

For any $\tau$,

$$\hat{g}_{K\tau} = 0 \iff \hat{g}_{A\tau} = 0. \qquad (84)$$

The "$\Rightarrow$" direction follows from (78). If $\hat{g}_{K\tau} = 0$, then $\sigma g_{A\tau} = (1-a)g_{K\tau}$; so if the right-hand side is constant around $\tau$, so is the left. The "$\Leftarrow$" direction follows likewise from (79).

Also, $\hat{g}_K$ and $\hat{g}_A$ are continuous in $t$ wherever they are defined. So by the intermediate value theorem, if either term is negative at some time and positive at another time, it must equal zero at an intermediate time. By (84), we must then have $\gamma = 1$.

It follows that, if $\gamma \neq 1$, either

1. $\hat{g}_{Kt} > 0$ and $\hat{g}_{At} > 0\ \forall t$,
2. $\hat{g}_{Kt} > 0$ and $\hat{g}_{At} < 0\ \forall t$,
3. $\hat{g}_{Kt} < 0$ and $\hat{g}_{At} > 0\ \forall t$, or
4. $\hat{g}_{Kt} < 0$ and $\hat{g}_{At} < 0\ \forall t$,

with case 4 incompatible with $\gamma > 1$ and case 1 incompatible with $\gamma < 1$. We will now show that cases 2 and 3 are also incompatible with $\gamma \neq 1$.

Consider case 2. From $\hat{g}_{Kt} > 0\ \forall t$, and (78), it follows that

$$g_{At} > (1-a)g_{Kt}\ \forall t. \qquad (85)$$

Recall that, by stipulation, $g_K$ always rising and $g_A$ is always falling. Thus $\{g_{Kt}\}$ is bounded above, for instance by $g_{A0}/(1-a)$, and $\{g_{At}\}$ is bounded below, for instance by $(1-a)g_{K0}$. By the monotone convergence theorem for functions, $\lim_{t\to\infty} g_{Kt}$ and $\lim_{t\to\infty} g_{At}$ are defined (and finite, and—since $g_{K0} = s(1-S_K)^a A_0 K_0^{a-1} > 0$—positive). Let us denote these limits $g_K^*$ and $g_A^*$ respectively.

By (78) and (79), it then follows that $\lim_{t\to\infty} \hat{g}_{Kt}$ and $\lim_{t\to\infty} \hat{g}_{At}$ are also defined (and finite). Since $g_K^*$ and $g_A^*$ are finite and nonzero, as we have just shown, it must be that $\lim_{t\to\infty} \hat{g}_{Kt} = \lim_{t\to\infty} \hat{g}_{At} = 0$. Taking the limits of terms (78) and (79), we then have

$$g_A^* = (1-a)g_K^* \text{ and} \tag{86}$$

$$g_K^* = \frac{1-\phi}{\lambda} g_A^*, \tag{87}$$

which jointly imply $g_A^* = \gamma g_A^*$ and thus $\gamma = 1$.

Case 3 can be shown to imply $\gamma = 1$ by a precisely analogous proof. Thus $\gamma > 1$ implies case 1 and $\gamma < 1$ implies case 4.

Suppose $\gamma > 1$. By the statements of case 1 and expressions (78)–(79), we have

$$g_{At} > (1-a)g_{Kt} \ \forall t \text{ and} \tag{88}$$

$$g_{Kt} > \frac{1-\phi}{\lambda} g_{At} \ \forall t. \tag{89}$$

By (88), and substituting by expressions (76) and (77),

$$\begin{aligned} g_{At}^2 &> (1-a)g_{At}\, g_{Kt} \\ &= \tilde{\theta} A_t^{\phi} K_t^{\lambda+a-1} \ \forall t, \end{aligned} \tag{90}$$

where

$$\tilde{\theta} \triangleq (1-a)s\theta(1-S_K)^a S_K^{\lambda}. \tag{91}$$

If the relationship of (88) were an equality at all $t$, then $A$ would always grow at precisely the same proportional rate as $K^{1-a}$. Noting that

$$A_0 = A_0 K_0^{a-1} \cdot K_0^{1-a}, \tag{92}$$

we would maintain this ratio between $A$ and $K^{1-a}$, with

$$A_t = A_0 K_0^{a-1} K_t^{1-a} \; \forall t. \tag{93}$$

It thus follows from (88) that

$$A_t \geq A_0 K_0^{a-1} K_t^{1-a} \; \forall t \tag{94}$$

$$\implies K_t \leq K_0 A_0^{a-1} A_t^{\frac{1}{1-a}} \; \forall t \tag{95}$$

(with equality at $t = 0$ and strict inequality at $t > 0$). It likewise follows from (89) that

$$K_t \geq K_0 A_0^{\frac{\phi-1}{\lambda}} A_t^{\frac{1-\phi}{\lambda}} \; \forall t. \tag{96}$$

So, if $\lambda + a - 1 \leq 0$, it follows from (90) and (95) that

$$g_{At}^2 > \tilde{\theta} A_0^{-\frac{\lambda+a-1}{1-a}} K_0^{\lambda+a-1} A_t^{\phi+\frac{\lambda+a-1}{1-a}} \; \forall t. \tag{97}$$

Given $\gamma > 1$, the exponent on $A_t$ in (97) is positive. Likewise, if $\lambda+a-1 > 0$, it follows from (90) and (96) that

$$g_{At}^2 > \tilde{\theta} A_0^{-\frac{1-\phi}{\lambda}(\lambda+a-1)} K_0^{\lambda+\alpha-1} A_t^{\phi+\frac{1-\phi}{\lambda}(\lambda+a-1)} \; \forall t. \tag{98}$$

Again, given $\gamma > 1$, the exponent on $A_t$ in (98) is positive. Either way, therefore, $A$ grows at worst hyperbolically, and so exhibits a Type II growth explosion. It follows immediately that $Y$ does as well.

If $\gamma < 1$, a proof that $A$ grows at best power-functionally is precisely analogous, except that it uses inequality (96) in the $\lambda + a - 1 \leq 0$ case and inequality (95) in the $\lambda + a - 1 > 0$ case. By (79) and the case 4 stipulation that $\hat{g}_{At} < 0 \; \forall t$, we then have

$$g_{Kt} < \frac{1-\phi}{\lambda} g_{At} \; \forall t, \tag{99}$$

implying that $K$ also grows at best power-functionally. Thus $Y$ grows at best power-functionally as well.

Furthermore, it follows from (75) that if $K$ were constant, $A$ (and thus $Y$) would grow power-functionally. Since the possibility of capital accumulation cannot decelerate output growth, $Y$ does in fact grow power-functionally.

Let us last consider the case of $\gamma = 1$.

From (78), we know that if $g_{Kt} > (<) \frac{1}{1-a} g_{At}$ then $\hat{g}_{Kt} < (>) 0$. Likewise, from (79), we know that if $g_{At} > (<) \frac{\beta}{1-\phi} g_{At}$ then $\hat{g}_{At} < (>) 0$. When $\gamma = 1$, however,

$$\frac{\beta}{1-\phi} = 1 - a, \tag{100}$$

so

$$\begin{align} & g_{K0} \geq g_{A0}/(1-a) \tag{101} \\ \iff & g_{Kt} \geq g_{At}/(1-a) \ \forall t. \tag{102} \end{align}$$

By the reasoning following (85), the limits $g_K^* \triangleq \lim_{t\to\infty} g_{Kt}$ and $g_A^* \triangleq \lim_{t\to\infty} g_{At}$ are defined, finite, and positive. Furthermore, by the continuity of $\hat{g}_K$ and $\hat{g}_A$ in $g_K$ and $g_A$, we must have $g_K^* = g_A^*/(1-a)$. Thus

$$\begin{align} \frac{g_K^*}{g_A^*} &= \frac{1}{1-a} \tag{103} \\ \Longrightarrow \lim_{t\to\infty} \frac{s(1-S_K)^a}{\theta S_K^\lambda} A_t^{2-\phi} K_t^{a-1-\lambda} &= \frac{1}{1-a} \tag{104} \\ \Longrightarrow \lim_{t\to\infty} A_t K_t^{a-1} = \left( \frac{\theta S_K^\lambda}{s(1-a)(1-S_K)^a} \right)^{\frac{1-a}{1+\lambda-a}} & \qquad \text{by } \gamma = 1 \tag{105} \\ \Longrightarrow g_K^* = s^{\frac{\lambda}{1+\lambda-a}} \left( \frac{\theta}{1-a} \frac{S_K^\lambda}{(1-S_K)^a} \right)^{\frac{1-a}{1+\lambda-a}} & \qquad \text{by (76)}. \tag{106} \end{align}$$

Finally,

$$\begin{align} \lim_{t\to\infty} g_{Yt} &= g_A^* + a g_K^* \tag{107} \\ &= g_K^* \qquad \text{by (103)}. \tag{108} \end{align}$$

# References


**Acemoglu, Daron**, *Introduction to Modern Economic Growth*, Princeton University Press, 2009.

_ **and David Autor**, "Skills, tasks and technologies: Implications for employment and earnings," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

_ **and** _ , "What Does Human Capital Do? A Review of Goldin and Katz's 'The Race Between Education and Technology'," *Journal of Economic Literature*, 2012, *50* (2), 426–463.

_ **and Pascual Restrepo**, "Modeling Automation," *AEA Papers and Proceedings*, 2018, *108*, 48–53.

_ **and** _ , "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 2018, *108* (6), 1488–1542.

_ **and** _ , "Artificial Intelligence, Automation and Work," in Ajay Agrawal, Joshua Gans, and Avi Goldfarb, eds., *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Research, 2019, pp. 197–236.

_ **and** _ , "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 2019, *33* (2), 3–30.

_ **and** _ , "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

_ **and** _ , "Tasks, Automation, and the Rise in US Wage Inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

_ **and Todd Lensman**, "Regulating Transformative Technologies," 2023. Working paper.

_ **, Claire Lelarge, and Pascual Restrepo**, "Competing with Robots: Firm-Level Evidence from France," *AEA Papers and Proceedings*, 2020, *110*, 383–88.

**Aghion, Philippe, Benjamin F Jones, and Charles I Jones**, "Artificial Intelligence and Economic Growth," in Ajay Agrawal, Joshua Gans, and Avi Goldfarb, eds., *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Research, 2019, pp. 237–282.

**Agrawal, Ajay, John McHale, and Alex Oettl**, "Finding Needles in Haystacks: Artificial Intelligence and Recombinant Growth," in Ajay Agrawal, Joshua Gans, and Avi Goldfarb, eds., *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Research, 2019, pp. 149–174.

_ **, Joshua Gans, and Avi Goldfarb**, *Prediction Machines: The Simple Economics of Artificial Intelligence*, Harvard Business Press, 2018.

**Aschenbrenner, Leopold**, "Existential Risk and Growth," 2020. Global Priorities Institute, Working Paper No. 6-2020.

**Autor, David**, "Why Are There Still So Many Jobs? The History and Future of Workplace Automation," *Journal of Economic Perspectives*, 2015, *29* (3), 3–30.

**Baily, Martin Neil, Erik Brynjolfsson, and Anton Korinek**, "Machines of mind: The case for an AI-powered productivity boom," *Brookings report*, 2023.

**Baumol, William J**, "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *The American Economic Review*, 1967, *57* (3), 415–426.

**Berg, Andrew, Edward F Buffie, and Luis-Felipe Zanna**, "Should We Fear the Robot Revolution? (The Correct Answer Is Yes)," *Journal of Monetary Economics*, 2018, *97*, 117–148.

**Besiroglu, Tamay, Nicholas Emery-Xu, and Neil Thompson**, "Economic impacts of AI-augmented R&D," *arXiv:2212.08198*, 2023.

**Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb**, "Are Ideas Getting Harder to Find?," *American Economic Review*, 2020, *110* (4), 1104–1144.

**Bolt, Jutta and Jan Luiten van Zanden**, "Maddison Project Database, version 2020," `https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2020` 2020.

**Bommasani, Rishi, Drew A. Hudson, Ehsan Adeli, Russ Altman et al.**, "On the Opportunities and Risks of Foundation Models," *arXiv:2108.07258*, 2021.

**Bostrom, Nick**, *Superintelligence*, Oxford University Press, 2014.

**Burks, Arthur W and John Von Neumann**, *Theory of Self-reproducing Automata*, University of Illinois Press, 1966.

**Carlsmith, Joseph**, "How Much Computational Power Does It Take to Match the Human Brain?," *Open Philanthropy Report*, 2020.

**Chirinko, Robert S and Debdulal Mallick**, "The Substitution Elasticity, Factor Shares, and the Low-frequency Panel Model," *American Economic Journal: Macroeconomics*, 2017, *9* (4), 225–253.

**Cockburn, Iain M, Rebecca Henderson, and Scott Stern**, "The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis," in Ajay Agrawal, Joshua Gans, and Avi Goldfarb, eds., *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Research, 2019, pp. 115–146.

**Erdil, Ege and Tamay Besiroglu**, "Explosive Growth from AI Automation: A Review of the Arguments," 2023. Working paper.

**Federico, Giovanni, Alessandro Nuvolari, and Michelangelo Vasta**, "Inequality in Pre-industrial Europe (1260-1850): New Evidence from the Labour Share," 2020. NYU Abu Dhabi, Working Paper No. 0051.

**Goldin, Claudia Dale and Lawrence F Katz**, *The Race Between Education and Technology*, Harvard University Press, 2009.

**Goldman Sachs**, "Generative AI could raise global GDP by 7%," `https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html` 2023.

**Gordon, Robert J**, *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*, Princeton University Press, 2016.

**Griliches, Zvi**, "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 1957, *25* (4), 501–522.

**Growiec, Jakub**, *Accelerating Economic Growth: Lessons From 200 000 Years of Technological Progress and Human Development*, Springer, 2022.

_ , "Automation, Partial and Full," *Macroeconomic Dynamics*, 2022, *26* (7), 1731–1755.

_ **, Julia Jabłońska, and Aleksandra Parteka**, "Hardware and Software over the Course of Long-Run Growth: Theory and Evidence," *SGH KAE Working Paper*, 2023, (2023/091).

**Hanson, Robin**, "Economic Growth Given Machine Intelligence," 2001. Working paper.

**Hémous, David and Morten Olsen**, "The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality," 2014. Centre for Economic Policy Research, Discussion Paper No. DP10244.

**Jones, Charles I**, "R&D-based Models of Economic Growth," *Journal of Political Economy*, 1995, *103* (4), 759–784.

— , "The A.I. Dilemma: Growth versus Existential Risk," *Working paper, Stanford*, 2023.

**Kaldor, Nicholas**, "A Model of Economic Growth," *The Economic Journal*, 1957, *67* (268), 591–624.

**Klinova, Katya and Anton Korinek**, "AI and Shared Prosperity," *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES 2021)*, 2021.

**Korinek, Anton**, "The Rise of Artificially Intelligent Agents," *Working paper, University of Virginia*, 2018.

— , "Generative AI for Economic Research: Use Cases and Implications for Economists," *Journal of Economic Literature*, 2023, *61* (4).

— **and Joseph E Stiglitz**, "Artificial Intelligence and Its Implications for Income Distribution and Unemployment," in Ajay Agrawal, Joshua Gans, and Avi Goldfarb, eds., *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Research, 2019, pp. 349–390.

— **and Megan Juelfs**, "Preparing for the (Non-Existent?) Future of Work," in Justin Bullock et al., ed., *forthcoming, Oxford Handbook of AI Governance*, Oxford University Press, 2023.

**Martin, Ian W. R. and Robert S. Pindyck**, "Averting Catastrophes: The Strange Economics of Scylla and Charybdis," October 2015, *105* (10), 2947–2985.

**Mookherjee, Dilip and Debraj Ray**, "Capital and Inequality in the Long Run: Automation Without Technical Progress," 2017. Working paper.

**Nordhaus, William D**, "Schumpeterian Profits in the American Economy: Theory and Measurement," Technical Report, National Bureau of Economic Research 2004.

— , "Are We Approaching an Economic Singularity? Information Technology and the Future of Economic Growth," *American Economic Journal: Macroeconomics*, 2021, *13* (1), 299–332.

**Oberfield, Ezra and Devesh Raval**, "Micro Data and Macro Technology," *Econometrica*, 2021, *89* (2), 703–732.

**Romer, Paul**, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–S102.

**Roodman, David**, "On the Probability Distribution of Long-term Changes in the Growth Rate of the Global Economy: An Outside View," 2020. Open Philanthropy.

**Sachs, Jeffrey D and Laurence J Kotlikoff**, "Smart Machines and Long-term Misery," Technical Report 2012. National Bureau of Economic Research, Working Paper No. 18629.

**_ , Seth G Benzell, and Guillermo LaGarda**, "Robots: Curse or Blessing? A Basic Framework," 2015. National Bureau of Economic Research, Working Paper No. 21091.

**Sandberg, Anders**, "An Overview of Models of Technological Singularity," in Max More and Natasha Vita-More, eds., *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, Wiley Online Library, 2013, pp. 376–394.

**Seeber, Isabella, Eva Bittner, Robert O Briggs, Triparna de Vreede, Gert-Jan De Vreede, Aaron Elkins, Ronald Maier, Alexander B Merz, Sarah Oeste-Reiß, Nils Randrup, Gerhard Schwabe, and Matthias Söllner**, "Machines as Teammates: A Research Agenda on AI in Team Collaboration," *Information & Management*, 2020, *57* (2), 103174.

**Sevilla, Jaime, Lennart Heim, Anson Ho, Tamay Besiroglu, Marius Hobbhahn, and Pablo Villalobos**, "Compute Trends Across Three Eras of Machine Learning," *arXiv:2202.05924*, 2022.

**Simon, Herbert A**, *The New Science of Management Decision*, Harper & Brothers, 1960.

**Uzawa, Hirofumi**, "Neutral Inventions and the Stability of Growth Equilibrium," *The Review of Economic Studies*, 1961, *28* (2), 117–124.

**Villalobos, Pablo**, "Scaling Laws Literature Review," 2023. Epoch AI blog post.

**Weitzman, Martin L**, "Recombinant Growth," *The Quarterly Journal of Economics*, 1998, *113* (2), 331–360.

**Yudkowsky, Eliezer**, "Intelligence Explosion Microeconomics," 2013. Machine Intelligence Research Institute, Technical Report.

**Zeira, Joseph**, "Workers, Machines, and Economic Growth," *Quarterly Journal of Economics*, 1998, *113* (4), 1091–1117.