THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

Dec/2019
Working Paper 19-18
rcea.org/RePEc/pdf/wp19-18.pdf

THE HARDWARE–SOFTWARE MODEL: A NEW CONCEPTUAL FRAMEWORK OF PRODUCTION, R&D, AND GROWTH WITH AI

Jakub Growiec
Warsaw School of Economics, Poland
RCEA

Rimini Centre for Economic Analysis
75 University Ave W., LH3079, Waterloo, Ontario, Canada, N2L3C5
www.rcea.org

The Hardware–Software Model: A New Conceptual Framework of Production, R&D, and Growth with AI*

Jakub Growiec†

December 2, 2019

Abstract

The article proposes a new conceptual framework for capturing production, R&D, and economic growth in aggregative economic models which extend their horizon into the digital era. Two key factors of production are considered: *hardware*, including physical labor, traditional physical capital and programmable hardware, and *software*, encompassing human cognitive work and pre-programmed software, including artificial intelligence (AI). Hardware and software are complementary in production whereas their constituent components are mutually substitutable. The framework generalizes, among others, the standard model of production with capital and labor, models with capital–skill complementarity and skill-biased technical change, and unified growth theories embracing also the pre-industrial period. It offers a clear conceptual distinction between mechanization and automation of production. It delivers sharp, empirically testable and economically intuitive predictions for long-run growth, the evolution of factor shares, and the direction of technical change.

Keywords: production function, R&D equation, technological progress, complementarity, automation, artificial intelligence.

JEL codes: O30, O40, O41.

*Thanks to Pedro Gil, Michał Lewandowski, Jakub Mućk, Klaus Prettner, Aleksandar Vasilev, Łukasz Woźny, conference participants at EEA/ESEM 2019 (Manchester), EAEPE 2019, RCEA and WIEM (Warsaw), and seminar participants in Warsaw, Gdańsk and Porto for helpful comments and suggestions. Financial support from the Polish National Science Center (Narodowe Centrum Nauki) under grant OPUS 14 No. 2017/27/B/HS4/00189 is gratefully acknowledged. All errors are my responsibility.

†Department of Quantitative Economics, SGH Warsaw School of Economics, Poland. Address: al. Niepodległości 162, 02-554 Warszawa, Poland. E-mail: jakub.growiec@sgh.waw.pl.

1 Introduction

The world economy has changed a lot since the 1980s. Pre-existing long-run trends in economic development like Kaldor's "stylized facts" (Kaldor, 1961) and the seemingly eternal constancy of "great ratios" (Klein and Kosobud, 1961) have been overturned, and new ones emerged (Jones and Romer, 2010). Among the new tendencies, during the last 40 years the world has been witnessing (even if only recently documenting) systematically declining labor shares (Arpaia, Pérez, and Pichelmann, 2009; Elsby, Hobijn, and Sahin, 2013; Karabarbounis and Neiman, 2014), increasing profit shares (Barkai, 2017), increasing markups and market power (De Loecker and Eeckhout, 2017, 2018; Diez, Leigh, and Tambunlertchai, 2018), increasing market concentration (Autor, Dorn, Katz, Patterson, and Van Reenen, 2017) and increasing within-country income inequality (Piketty, 2014; Piketty and Zucman, 2014; Milanović, 2016). All this was accompanied by a tendency of skill polarization, gradual elimination of routine jobs (Acemoglu and Autor, 2011; Autor and Dorn, 2013), and an increasing variety of jobs becoming susceptible to automation (Frey and Osborne, 2017; Arntz, Gregory, and Zierahn, 2016).

By contrast, our workhorse economic growth models (Barro and Sala-i-Martin, 2003; Jones, 2005a; Acemoglu, 2009) imply stable factor shares, markups and market concentration over the long run, stationary income inequality and a fixed steady-state job structure. They are therefore unable to reconcile the pre-1980 growth experience with the emerging new regularities. Established unified growth theories (Galor and Weil, 2000; Galor, 2005, 2011), despite successfully explaining the mechanisms of transition from a stagnant agricultural to a growing industrial economy during the Industrial Revolution, tend to be equally ill-suited to capturing the new tendencies. Looking through the lens of the conventional growth theories, one cannot help but classify the new global macro trends as *puzzles*.

A likely reason for the apparent mismatch between data and theory is that except for a few forerunners (Acemoglu and Restrepo, 2018; Benzell, Kotlikoff, LaGarda, and Sachs, 2015; Berg, Buffie, and Zanna, 2018), growth models developed thus far have been either rooted entirely in the industrial era, or focused on even earlier eras. They generally do not acknowledge that since the 1980s the Digital Revolution has been transforming the world before our eyes in a comparably profound way to what the Industrial Revolution had done two centuries ago. The computer age – to kindly paraphrase Robert Solow – is now seen everywhere, even in productivity statistics. Since the 1980s personal computers have been permeating firms and households, and digitization gained massive momentum in the 2000s with the spread of the Internet. Quantitatively, since the 1980s "general-purpose computing capacity grew at an annual rate of 58%. The world's capacity for bidirectional telecommunication grew at 28% per year, closely followed by the increase in globally stored information (23%)" (Hilbert and López, 2011). The costs of a standard computation have been declining by 53% per year on average since 1940 (Nordhaus, 2017). Hence, growth in the digital sphere is now an order of magnitude faster than growth in the global capital stock and GDP: data volume, processing power and bandwidth double every 2–3 years, whereas global GDP doubles every 30–35 years. The processing, storage, and communication of information has decoupled from the cognitive capacities of the human brain; "less than one percent of information was in digital format in the mid-1980s, growing to more than 99% today" (Gillings, Hilbert, and Kemp, 2016). Preliminary evidence also suggests that since the 1980s the efficiency of computer algorithms has been improving at a pace that is of the same order of magnitude as accumulation of digital hardware (Grace, 2013). Corroborating this finding, in the recent decade we have witnessed a surge in AI breakthroughs based on the methodology of *deep neural networks* (Tegmark, 2017), from autonomous vehicles and simultaneous language interpretation to self-taught superhuman performance at chess and Go (Silver, Hubert, Schrittwieser, et al., 2018). We are also observing that ever since Bill Gates first topped the list of World's Billionaires in 1995, biggest fortunes these days are made in the computer software business.

The objective of the current paper is to propose a new conceptual framework for modeling long-run economic growth, compatible both with pre-1980 macro trends and the present world where information processing, communication and storage is increasingly detached from human minds. To this end I take a big step back and, with all the emerging new macro and technological trends in mind, re-evaluate the key inputs to aggregate production and R&D. I find that the classical capital–labor dichotomy, on which virtually all existing models are based, does not sufficiently describe the supply side of the digital-era economy which features also pre-programmed software (including AI algorithms), able to operate without any human input. My proposition is to replace capital and labor as key factors of production with broader aggregates which I call *hardware* and *software*. Based on these new concepts I lay out the rudiments of a macroeconomic framework for modeling production, R&D and growth across the human history, including and specially focusing on the digital era. I demonstrate that the new framework, the *hardware–software model*, allows to adapt our existing growth models to the realities of the incipient digital era without sacrificing the accuracy in describing the past.

The key premise of the proposed new framework lies with the postulate that valuable output can only be generated through purposefully initiated physical action. Thus, generating output (either in the material or in the informational form) requires both some physical *action* and some *code*, a set of instructions describing the action. In consequence, the general form of any production function should feature some physical *hardware* X, performing the action, and some disembodied *software* S, providing the relevant information. This simple observation has profound consequences. It underscores that physical capital and human physical labor should be modeled as fundamentally substitutable inputs, contributing to *hardware*: they are both means by which we perform physical action. Analogously, human cognitive work and pre-programmed software should also be viewed as substitutes, making up the *software* factor: they are the source of instructions for the performed action. In turn, both hardware and software are clearly complementary and essential in the process. (By *complementarity*, in practice I will mean their gross complementarity in the sense of elasticity of substitution being below unity.) Furthermore, programmable hardware, such as computers, smartphones or robots, similarly to the human body has double duty: as means of performing physical actions and as a container for software – stored information and working algorithms.

The hardware–software model delivers a range of testable predictions, allowing for its empirical assessment and potential falsification. First, it predicts that soon after an industrial revolution – understood as emergence of new accumulable components of *hardware* – physical labor should be gradually replaced. The physical labor share of output should then go down, and rents to capital, energy and cognitive work should go up. The overall labor share of output should first go down and then up. Analogously, after a digital revolution – understood as emergence of new accumulable components of *software* – cognitive jobs should be gradually automated. The cognitive and overall labor share should then go down, whereas rents to pre-programmed software, data and programmable hardware should go up. The physical capital share of output should first go down and then up. These predictions appear to be (at least qualitatively) consistent with the available evidence, especially when admitting that software and data rents are generally not separately accounted and enter into firms' profits. Second, the proposed framework re-structures growth accounting, implying in particular that accumulation of programmable hardware and increases in working population contribute both to hardware and software, with specific time-varying shares. It also provides a parsimonious testable prediction that all technical change should be *software-augmenting*. Third, the framework imposes testable restrictions on production functions both for aggregate output and for ideas (the R&D equation). For example, it expects that R&D workers should be complementary to lab equipment (R&D hardware) in producing R&D output.

The hardware–software model is helpful in explaining global long-run growth processes also because it nests the following conventional models as special cases:

(i) a standard treatment of an industrial economy producing with capital and labor and respecting Kaldor's facts (obtained by assuming that all physical work is done by machines and all cognitive work is done by humans),

(ii) a model of capital–skill complementarity and skill-biased technical change (assuming that all cognitive work is done by humans),

(iii) a unified growth theory addressing the period of Industrial Revolution (after the arrival of new accumulable hardware),

(iv) a theory of inception and further development of the digital era (after the arrival of new accumulable software).

In the policy perspective, the hardware–software model informs the debate on the future of global economic growth – whether we should expect secular stagnation (Jones, 2002; Gordon, 2016), balanced growth with limited automation, "race against the machine" (Acemoglu and Restrepo, 2018), or technological singularity (Kurzweil, 2005). It organizes the predictions for the global future in the following way. First, in the digital era, as production gets increasingly automated, software gradually decouples from human cognitive work and becomes proportional to hardware because pre-programmed software can be virtually costlessly copied and thus can easily scale up to the level of available programmable hardware. Under constant returns to scale and in the absence of further technological revolutions[1], this gradually reduces the role of skill-biased technical change and eventually generates long-run endogenous growth by hardware accumulation alone. In the limit, all production is automated. Second, complementarity and substitutability shape the dynamics of factor shares and global inequality. The Industrial Revolution had vastly different implications for factor shares than the ongoing Digital Revolution because the former featured replacement of humans with machines in the hardware factor (brawn) whereas the latter pertains to the software factor (brains). The Industrial Revolution (or the process of mechanization) raised demand for human skilled labor; the Digital Revolution (or the process of automation) replaces human skilled labor and raises demand only for complementary computer hardware.

This paper is related to at least five strands of literature. First, the literature on production function specification and estimation, in particular with capital–skill complementarity, unbalanced growth, as well as investment-specific and skill-biased technical change.[2] Second, the literature preoccupied with accounting for the accumulation of information and communication technologies (ICT) and their broad growth-enhancing role as a general purpose technology.[3] Third, studies focusing on automation and its impacts on productivity, employment, wages and factor shares.[4]

[1]Given the observed pace of growth in computing power and AI capabilities, further technological revolutions are actually quite likely, though.

[2]Including among others Gordon (1990); Jorgenson (1995); Greenwood, Hercowitz, and Krusell (1997); Hercowitz (1998); Kumar and Russell (2002); Koop, Osiewalski, and Steel (1999, 2000); Krusell, Ohanian, Ríos-Rull, and Violante (2000); Henderson and Russell (2005); Caselli and Coleman (2006); Klump, McAdam, and Willman (2007, 2012); Growiec (2012); Mućk (2017); McAdam and Willman (2018).

[3]Including among others Bresnahan and Trajtenberg (1995); Timmer and van Ark (2005); Jorgenson (2005); Brynjolfsson and McAfee (2014); Gordon (2016); Brynjolfsson, Rock, and Syverson (2019); Nordhaus (2017); Aum, Lee, and Shin (2018).

[4]Including among others Acemoglu and Autor (2011); Autor and Dorn (2013); Graetz and

Fourth, the nascent literature on macroeconomic implications of development of AI and autonomous robots.[5] Last but not least, the voluminous literature on R&D based endogenous growth.[6]

The remainder of the paper is structured as follows. Section 2 defines the factors of production of the hardware–software model. Section 3 discusses the conceptual underpinnings of the aggregate production function. Section 4 tackles the R&D equation. Section 5 discusses the properties of a fully specified growth model à la Solow (1956); Mankiw, Romer, and Weil (1992) with CES production. Section 6 concludes with a general discussion of the framework, spelling out the key concepts and misconceptions of the digital era, and speculating about the future.

2 The Hardware–Software Model

In any conceivable technological process, valuable output is generated through some physical action. It is a local reduction of entropy, and so it typically does not occur by chance but is purposefully initiated. In other words, producing valuable output requires both some physical *action* and some *code*, a set of instructions describing the action. Based on this premise I posit that the general production function (for whatever output) should feature some physical *hardware* X, able to perform the action, and some disembodied *software* S, providing information on what should be done and how. This naturally leads to a general form:

$$Output = \mathcal{F}(X, S), \tag{1}$$

where $\mathcal{F}$ is increasing and concave in both factors and such that hardware X and software S are essential and mutually complementary. The degree of their complementarity is an open question; the plausible range spans from perfect complementarity (Leontief form) if just one method of producing output exists, to imperfect complementarity if producers are allowed to choose their preferred technology from a technology menu (Jones, 2005b; Growiec, 2013, 2018). Intuitively, a little substitutability is likely because the same outcome can sometimes be generated with more resources (larger X) but less efficient code (smaller S), or vice versa, but the fundamental complementarity should prevail. One natural way to instantiate this

Michaels (2018); Acemoglu and Restrepo (2018); Andrews, Criscuolo, and Gal (2016); Arntz, Gregory, and Zierahn (2016); Frey and Osborne (2017); Barkai (2017); Autor, Dorn, Katz, Patterson, and Van Reenen (2017); Jones and Kim (2018); Hemous and Olsen (2018).

[5] Including Yudkowsky (2013); Graetz and Michaels (2018); Sachs, Benzell, and LaGarda (2015); Benzell, Kotlikoff, LaGarda, and Sachs (2015); DeCanio (2016); Acemoglu and Restrepo (2018); Aghion, Jones, and Jones (2019); Berg, Buffie, and Zanna (2018).

[6] Including among others Romer (1990); Jones and Manuelli (1990); Aghion and Howitt (1992); Jones (1995); Acemoglu (2003); Ha and Howitt (2007); Madsen (2008); Bloom, Jones, Van Reenen, and Webb (2017); Kruse-Andersen (2017).

assumption is to take a CES specification with an elasticity of substitution $\sigma \in (0,1)$, cf. Klump, McAdam, and Willman (2007, 2012). The particular CES form of the $\mathcal{F}$ function is however not necessary for the results.[7]

The specification (1) abstracts from raw materials, energy and data which are being used up in the production process. It works as if we assumed that they were given for free and in infinite supply, or at least that they were sufficiently cheap and abundant that they would never become a bottleneck. Relaxing this simplifying assumption is left for further research.

Hardware X encompasses physical actions performed by both humans and machines. Hence, X encompasses both the services of physical capital K and of physical labor L, where the latter variable excludes any know-how or skill of the worker.

Software S, in turn, encompasses all useful instructions which stem from the available information, in particular the practical implementation of state-of-the-art technologies. Hence, it includes the skills and technological knowledge employed in human cognitive work, H, as well as pre-programmed software Ψ, which is essentially a task-specific list of instructions to be performed by the associated programmable hardware.[8] Software Ψ may in particular include artificial intelligence (AI) algorithms, defined as the software which is able to learn from data as well as potentially self-improve and self-replicate. I implicitly assume that there are no physical obstacles precluding pre-programmed software from performing (or more precisely, providing the hardware with instructions to perform) any task available to a human (Yudkowsky, 2013; Dennett, 2017).

Within hardware, I view the agents of physical action as perfectly substitutable. This reflects the idea that whatever it is that performs a given set of actions, if the actions are the same then the outcome should be the same, too. The same logic applies to software: regardless of whether a set of instructions comes from a human brain or a mechanical information processing unit, if the actual information content of instructions is the same, then the outcome should be the same, too.[9]

[7]For example, Growiec and Mućk (2018) propose a more flexible parametric framework that also allows the modeler to control whether the factors are gross substitutes or gross complements.

[8]Contemporary programmable hardware consists typically of computers, robots, and other devices embodying digital chips. In principle, it does not have be silicon-based, though. In fact the first pieces of non-biological programmable hardware were mechanical devices such as the Jacquard loom using punchcards, first invented in 1804.

[9]An important caveat is that by saying this I exclude complex, multi-step tasks that have not been yet fully automated. For example, if a cognitive task consists of two necessary steps, the first of which can be performed by a computer algorithm but the latter (under current technology) only by a human, then pre-programmed software and human cognitive work will turn out complementary at the level of the whole task even though they are perfectly substitutable at the level of the two sub-tasks. This apparent complementarity disappears, however, once the whole task becomes fully automatable. A more detailed treatment of complex tasks within the hardware–software model is an important objective for further research. It should also be said that technically speaking *perfect* substitutability is not necessary for the key results. At the cost of less

This leads to the following formula:

$$Output = \mathcal{F}(X, S) = \mathcal{F}(L + K, H + \Psi). \tag{2}$$

Each of the four identified factors of production has its unique properties (Table 1).

- *Human physical labor* L is rivalrous and given in fixed supply per worker and unit of time, $L = \zeta N$ where $\zeta \in [0, \bar{\zeta}]$ denotes the supply of physical labor per worker in a unit of time, expressed in physical capital units, and N is the total number of workers.

- *Physical capital* K is rivalrous but can be unboundedly accumulated in per-capita terms. Physical capital K may be non-programmable or programmable. The share of programmable hardware in total physical capital is denoted by χ (so that $\chi \in [0, 1]$).

- *Human cognitive work* H consists of three components, technological knowledge A, the average skill level h, and the number of workers N, as in $H = AhN$. Technological knowledge A, or the size of the "repository of codes" is non-rivalrous (Romer, 1986, 1990) and accumulable.[10] Per-capita skill levels h are rivalrous and bounded above, theoretically by the optimal code for performing a given task, but in practice by a much lower number $\bar{h} > 0$ due to the human inability to rewire our brains in order to perform cognitive tasks more efficiently (Yudkowsky, 2013) and more down-to-earth reasons like human mortality and decreasing returns in education.

- *Pre-programmed software* Ψ also consists of three components, technological knowledge A, "AI skill level" ψ which captures the degree to which pre-programmed software is able to perform the tasks collected in A, and the stock of programmable hardware χK on which the software is run, as in $\Psi = A\psi\chi K$. Technological knowledge A is the same as above.[11] The AI skill level ψ is assumed to be bounded above by the optimal code for performing a given task (e.g., perfect accuracy), though there may be in fact a much lower upper bound $\bar{\psi}$ (Hanson and Yudkowsky, 2013).[12] Because software can be virtually cost-

transparent notation, one can straightforwardly generalize the hardware–software model to accomodate imperfect substitutability between people and machines in both hardware and software, as in $Output = \mathcal{F}(G_1(L, K), G_2(H, \Psi))$, with gross substitutability of factors within G_1 and G_2.

[10] Depending on the institutional setup (e.g., intellectual property rights), technological knowledge A may be characterized by varying levels of excludability.

[11] If in reality the sets of codes available to humans and AI are different, the discrepancy between the measures of both sets can be captured by the factor ψ relative to h.

[12] Depending on the institutional setup (e.g., proprietary code vs. open source), AI skill level ψ may be characterized by varying levels of excludability.

lessly copied, it is assumed that it can scale up to the level of all available programmable hardware χK.[13]

Table 1: Factors of Production and R&D

Hardware X	Human physical labor	$L = \zeta N$
	Non-programmable physical capital	$(1-\chi)K$
	Programmable physical capital	χK
Software S	Human cognitive work	$H = AhN$
	Pre-programmed software†	$\Psi = A\psi\chi K$

Note: † includes AI algorithms.

It is important to observe that the hardware–software model envisages technological progress (growth in A) as expansion of the "repository of codes", i.e., as the development of new, better instructions allowing to produce higher output with a given amount of hardware. These instructions can be applied to the tasks at hand both by humans and machines. This is intuitive: technological progress may take the form of new abstract ideas, scientific theories, systematically catalogued facts, codes specifying certain actions, or blueprints of physical items; all this is *information* and not actual *objects* or *actions*, and it is precisely this informational character that makes technologies non-rivalrous (Romer, 1986, 1990). Thus all technological progress is naturally modeled as *software-augmenting*. In the hardware–software model, in contrast to the standard capital–labor model, there is no room for discussion on the direction of technical change – a parsimonious property that is highly valuable from a reductionist point of view.

3 The Aggregate Production Function

The aggregate production function is a key element of any macroeconomic model, and particularly so of any long-run growth theory. Since the 1950s (Solow, 1956, 1957), it has become commonplace to take capital K and labor L as the key inputs of this function, and value added (or GDP) as its output Y. Furthermore, it has become a very frequent, if not default, practice to assume purely labor-augmenting (Harrod-neutral) technical change, as in

$$Y = F(K, AL). \tag{3}$$

Of course, like any aggregate production function specification (Temple, 2006), equation (3) is a simplification that disregards the fact that K and L are amalgamates of heterogeneous components. The key question is, though, whether this

[13]Which implies that, in its basic form, the model abstracts from economic and legal constraints on the diffusion of software, such as the protection of intellectual property rights.

simplified form is sufficient for capturing the key macroeconomic facts it is meant to represent. Unfortunately, evidence is mounting that it is actually not the case. From the literature[14] it is gradually becoming clear that the standard treatment of inputs as in (3) may have been sufficient to model the classic Kaldor (1961) facts but fails at capturing the new development facts which emphasize irreducible heterogeneity within the K and L factors. It also cannot capture the new phenomena specific to the digital era.

3.1 Setup

The hardware–software production function proposed in this paper, following directly from equation (2), specifies the production factors differently. It generalizes equation (3) in a way that allows for consistency both with the key historical macro facts and the incipient digital production technology using also pre-programmed software, including AI.

Using the concepts from the previous section, the aggregate production function F is formalized as:

$$Y = F(X, S) = F(\zeta N + K, A(hN + \psi\chi K)), \tag{4}$$

where Y is aggregate value added (or GDP). The function F is increasing and concave in both its arguments. The standard replication argument applied to this production function specification implies constant returns to scale with respect to the rivalrous factors X and $S/A = hN + \psi\chi K$. With respect to X, S/A and A, though, returns to scale are increasing (Romer, 1986, 1990).

From the laws of thermodynamics, implying in particular that performing physical action requires expediting energy, it is expected that an essential fraction of GDP must consist of material outputs, serving – at the very least – to sustain the hardware (including human bodies) and allow it to work (Georgescu-Roegen, 1971, 1975). This observation reinforces the assumption that hardware X must be essential in the production process.

Pre-programmed software can be deployed in production processes only if the technology allows for the existence of programmable hardware ($\chi > 0$). Once it is introduced, though, there is no upper bound for its capacity relative to the cognitive capacity of the human brain. It may even one day come to exhibit superhuman cognitive performance. This is because (i) the human brain has fixed computational capacity whereas pre-programmed software (including AI) can be run on programmable hardware with any level of computing power, (ii) AI algorithms have

[14] Such as Gordon (1990); Greenwood, Hercowitz, and Krusell (1997); Krusell, Ohanian, Ríos-Rull, and Violante (2000); Caselli and Coleman (2006); Klump, McAdam, and Willman (2007); Jones and Romer (2010); Growiec (2012); McAdam and Willman (2018).

the ability to learn from data and potentially self-improve their architecture. Nevertheless, even without superhuman AI performance all cognitive tasks are amenable to automation with sufficient computing power χK (see the discussion in Section 6). The only pre-condition for this outcome is that in the full model (such as e.g. the one presented in Section 5) the possibility of accumulating the requisite computing power is not precluded by, e.g., preferences or institutions.[15]

Equations (2) and (4) signify also that AI is viewed here just as (improved) computer software and not as a separate production input. This is because AI algorithms provide drastic improvements in the applicability, efficiency, and versatility of software, but do not constitute a qualitative change in its function as means of providing instructions to programmable hardware. Hence, the model does not feature a separate AI revolution, and rather sees AI development as a massive boost to the Digital Revolution which already began with the early computer hardware and software. In my view, AI is to the digital era what the development of electricity and internal combustion engines was to the industrial era: a second wave of key breakthroughs, forcefully accelerating the impact of the initial revolutionary technological ideas on the economy and society, but not a separate technological revolution (Gordon, 2016).

It is instructive to consider four special cases of the model, representing four distinct conventional frameworks.

Industrial economy producing with capital and labor. Under the assumption that all physical work is done by machines ($\zeta = 0$) and all cognitive work is done by humans ($\chi = 0$), the production function (4) reduces to the conventional capital–labor specification with purely labor-augmenting technichal change, $Y = F(K, AhN)$. Capital and labor are then naturally gross complements, as suggested by bulk of the recent empirical literature (Klump, McAdam, and Willman, 2007, 2012; Mućk, 2017).

Capital–skill complementarity and skill-biased technical change. Under the assumption that all cognitive work is done by humans ($\chi = 0$), the production function (4) reduces to the specification with capital-skill complementarity (Krusell, Ohanian, Ríos-Rull, and Violante, 2000; Caselli and Coleman, 2006; McAdam and Willman, 2018) and skill-biased (or more precisely, cognitive labor-augmenting) technical change, $Y = F(\zeta N + K, AhN)$. Gross complementarity between hardware and software implies that physical capital is complementary to cognitive ($\approx$ skilled) labor H but substitutable with physical ($\approx$ unskilled) labor L, in line with findings of the literature.

[15] However, in a more general model with complex, multi-step tasks, human cognitive work can become essential for generating output if at least one step of at least one essential task is not automatable. Essentiality implies that there is no way around this particular step and no possibility of substituting out the entire task.

Industrial Revolution. The hardware–software model represents the Industrial Revolution as an episode where physical capital begins to be accumulated after the initial restriction $K \approx 0$ is lifted. In result human physical labor is gradually replaced with machines within hardware, in a process which we may call *mechanization* of production.

Digital Revolution. The model represents the Digital Revolution as an episode where pre-programmed software begins to be accumulated after the initial restriction $\chi = 0$ (and thus $\Psi = 0$) is lifted. In result human cognitive work is gradually replaced with machine code within software, in a process which we may call *automation* of production.

3.2 Growth Accounting

Log-differentiating equation (4) with respect to time, I obtain the following Solow-type decomposition of economic growth:

$$g_Y = \pi_X g_X + \pi_S g_S, \tag{5}$$

where $\pi_X = \frac{\partial Y}{\partial X}\frac{X}{Y}$ is the hardware share of output, and analogously $\pi_S = \frac{\partial Y}{\partial S}\frac{S}{Y}$ is the software share. Due to constant returns to scale with respect to rivalrous inputs and purely software-augmenting technical change, $\pi_X + \pi_S = 1$.

Decomposing (4) further, I obtain:

$$g_Y = \pi_X[\pi_L g_N + \pi_K g_K] + \pi_S[\pi_H(g_h + g_N) + \pi_\Psi(g_\psi + g_\chi + g_K)] + \pi_S g_A, \tag{6}$$

where – due to the assumption of perfect substitutability of the constituent components of hardware and software – the shares are simply $\pi_L = \frac{L}{X}$, $\pi_K = \frac{K}{X}$, $\pi_H = \frac{H}{S}$ and $\pi_\Psi = \frac{\Psi}{S}$.

Equation (6) presents formally that there are multiple potential sources of output growth in the hardware–software model. Each of them has different asymptotic properties.

- Population growth g_N increases the total amounts of both human physical and cognitive work. If there is also physical capital or pre-programmed software in the economy, this impact is less than proportional to output growth and thus, *ceteris paribus*, growth in output per worker $(g_Y - g_N)$ decreases with population growth.

- Physical capital accumulation g_K affects output growth both directly via the hardware component and indirectly via the pre-programmed software component (if $\pi_\Psi > 0$). It is subject to decreasing returns, but to a decreasing degree, and as $\pi_K \to 1$ and $\pi_\Psi \to 1$ the returns become asymptotically constant, allowing for unbounded output growth (Jones and Manuelli, 1990).

- Growth in average skill level per worker g_h and in the AI skill level g_ψ can be decisive in the short to medium run, but their impact on growth is by definition transitory and bound to disappear as $h \to \bar{h}$ and $\psi \to \bar{\psi}$.
- Growth in the share of programmable hardware g_χ can be important in the short to medium run, but should not play any role over the long run because χ is bounded between zero and one.
- Technological change g_A, understood as the increase in technological knowledge A, the size of the "repository of codes", is conceptually independent of human and AI skill accumulation. It adds to output growth with an elasticity equal to the software share and can be potentially unbounded.

While the software-augmenting character of technological change comes out as a very natural implication of the proposed conceptual framework, it stands in stark contrast to the discussions in the literature on the direction of factor-augmenting technical change (e.g. Acemoglu, 2003; Jones, 2005b; León-Ledesma, McAdam, and Willman, 2010). This is because conventional production factors such as capital and labor conflate hardware and software. If in fact technical change augments *software*, though, then it runs orthogonal to the classic capital–labor divide: it affects cognitive work but not physical labor, and pre-programmed software but not the hardware on which it is run.

The new framework also resolves the conundrum whether technological progress is disembodied or embodied in new investment goods (e.g. Gordon, 1990; Greenwood, Hercowitz, and Krusell, 1997; Hercowitz, 1998): in itself, it is the *disembodied information* that allows for more efficient actions. Nevertheless it may require investment in the complementary hardware in order to deliver the desired effects for output.

3.3 Stages of Economic Development

Let us now trace how the hardware–software model can be used to capture the key properties of production processes across the human history.

Stage 1. Pre-industrial production. In a pre-industrial economy, output was produced primarily in farming. At that stage of development, there was no significant accumulation of productive capital K per capita. Output was produced with a technology that used only human (and animal) physical labor for performing the physical actions and required also the services of land, a vital but essentially fixed[16]

[16]By making this assumption I concentrate on a mature agricultural economy and exclude the periods of transition from hunting and gathering to sedentary agriculture or conquests of new agricultural land.

factor of agricultural production. There was also no pre-programmed software Ψ. Setting $K = \tilde{K}$, representing land, and $\chi = 0$ in equation (4) yields the following simple formula:

$$Y = F(X, S) = F(\zeta N + \tilde{K}, AhN) \approx N \cdot F(\zeta, Ah), \tag{7}$$

where the last approximation follows from the assumption that $\tilde{K}$ is fixed and small relative to ζN. Hence, under gross complementarity of hardware and software (actions and instructions), pre-industrial output per worker was bounded above due to the scarcity of hardware.

Stage 2. Industrial production. Following the Industrial Revolution (≈1800 CE onwards) human (and animal) physical labor was gradually replaced with machines in a process of *mechanization* of production. The stock of physical capital per worker K/N began to grow exponentially. Productive physical actions were, however, still dependent solely on the instructions produced through human cognitive work; there was no programmable hardware and no pre-programmed software yet. As hardware was accumulated faster than software, the latter eventually became relatively scarce, at which point demand for human cognitive skills began to grow, setting up a secular upward trend in wages (Galor, 2005). Setting $\chi = 0$ in (4) yields:

$$Y = F(X, S) = F(\zeta N + K, AhN). \tag{8}$$

The limit of full mechanization and skill satiation, $K \to \infty$ and $h \to \bar{h}$, where $\bar{h}$ is the upper limit of human capital (skill) accumulation, implies $Y = F(K, A\bar{h}N)$. Hence, under this specification we obtain – in the limit – the standard balanced growth path result with gross complementarity of inputs and purely "labor-augmenting" (though really software-augmenting) technical change (Uzawa, 1961; Acemoglu, 2003). Along the balanced growth path, K/N grows at the same rate as technological knowledge A.

Stage 3. Digital production. Following the Digital Revolution (≈1980 CE onwards) we are observing gradual *automation* of production. Human cognitive skills which scale with the working population N are replaced with pre-programmed routines which scale with programmable hardware χK that grows exponentially faster. Consequently, software-augmenting technical change no longer affects only the efficiency of human cognitive work, but also to an increasing degree the capacities of pre-programmed software. As automation progresses, skill-biased technical change gradually morphs into routine-biased technical change (Acemoglu and Autor, 2011; Autor and Dorn, 2013). This is the world in which we live now.

At a later stage of the digital era, however, conventional case-based software will likely be replaced with self-improving AI algorithms, allowing for multiple-fold increases in ψ (Berg, Buffie, and Zanna, 2018) and thus fortifying the emerging upward trend in the share of non-human component in software.

The limit of full automation implies

$$Y = K \cdot F(1, A\bar{\psi}\bar{\chi}), \tag{9}$$

where $\bar{\psi}$ is the upper limit of AI skill accumulation and $\bar{\chi} \in (0, 1]$ is the limiting share of programmable hardware in all physical capital as $K \to \infty$. Full automation of the production process in the limit means leaving no jobs in the production sector to be performed by humans or at least rendering their contribution to output negligibly small.[17]

Equation (9) delivers an AK-type implication: there is long-run endogenous growth due to the accumulation of (programmable) hardware alone (Jones and Manuelli, 1990; Barro and Sala-i-Martin, 2003). This striking result is driven by two forces: (i) that pre-programmed software expands proportionally with programmable hardware, and (ii) that hardware and software are gross complements, and thus in the long run the pace of hardware accumulation determines the pace of economic growth.

Although asymptotically constant, the pace of hardware accumulation (and thus economic growth under full automation) may be nevertheless stupefying, with doubling times of the order of 2–3 years. In contrast to this prediction, what has been bringing global economic growth down in the recent decades, was the large share of "traditional" (non-programmable) capital, and – crucially – lack of AI algorithms able to fully tap the available computing power. Neither of these two constraints is guaranteed to persist into the indefinite future, though.

Hypothetical stage 4. Post-digital production. Under high to full automation of production processes, programmable hardware χK will gradually become the bottleneck of further development, the key factor constraining its pace. This will increase the incentives to invest in R&D directed towards radical innovations holding the promise to eliminate this bottleneck.[18]

Formally, such an episode of "new mechanization" may be modelled by intro-

[17] Putting it more harshly, under full mechanization and automation human work becomes *useless* for the economy (Harari, 2017).

[18] Such breakthrough technology would have to tap an entirely new source of energy, fundamentally increase energy efficiency, or otherwise massively improve unit productivity of programmable hardware. Among the probable scenarios, one could envision the arrival of quantum computing (in which case the Google AI Quantum team has already achieved a major breakthrough, Arute, Arya, Babbush, et al. (2019)), disruptive nanotechnology, massively improved solar power cells, or perhaps something yet unimagined. Extrapolating past trends in information processing and data accumulation and expecting them to feed into R&D productivity (see the next section of this paper), it is conceivable that such new breakthrough technology may in fact arrive quite soon.

ducing an additional component to the hardware amalgamate, as in:

$$X = \zeta N + K + \omega M, \tag{10}$$

where M denotes the new form of hardware, and $\omega \gg 1$ captures its unit productivity relative to K. This form of hardware must be programmable, so that AI could scale with M and avoid becoming a growth bottleneck itself.

Long-run implications include gradual replacement of K-type hardware with M and a permanent acceleration in growth. In fact, this additional acceleration in hardware X accumulation may eventually lead to a new growth regime "with a doubling time measured in days, not years" (Hanson, 2000).

In a world with fully mechanized and automated production, a new form of programmable hardware M, and AI that is able to scale with M, in the limit of $K/M \to 0$ the aggregate production function becomes again linear:

$$Y = F(\omega M, A\bar{\psi}M) = M \cdot F(\omega, A\bar{\psi}). \tag{11}$$

This means that despite the new breakthrough and the acceleration, hardware remains the bottleneck (i.e., key factor constraining the pace) of long-run growth.

3.4 Factor Shares

The assumption of gross complementarity of production inputs (as exemplified by CES technology with $\sigma \in (0, 1)$) provides a clear-cut implication for factor shares: factor income will be disproportionately directed towards the scarce factor. The hardware–software model delivers the following (empirically testable and intuitively explicable) predictions.

Stage 1. Pre-industrial production. In a mature pre-industrial economy able to achieve systematic technological progress (growth in A), increasing scarcity of human physical labor and agricultural land ($\zeta N + \tilde{K}$) relative to human cognitive work (AhN) implies that an ever increasing portion of value added is directed to hardware at the expense of software. The counterfactual limit of $A \to \infty$ without an industrial revolution (with a fixed $K = \tilde{K}$) implies a zero software share of output as virtually all revenues are directed towards scarce "hardware" which precludes further growth in per capita output: agricultural land and agricultural workers.

Stage 2. Industrial production. The first stage of development of an industrial economy features gradual *mechanization* of production: physical capital accumulation systematically reduces the role of human physical labor in hardware. Given the substitutability between capital K and physical labor ζN, the physical labor share goes down whereas the capital share goes up – a trend which Karl Marx called "the exploitation of the working class".

However, if the pace of capital accumulation during the Industrial Revolution outruns technical change (growth in A), this secular trend is accompanied also by an increasing output share accruing to software (i.e., human cognitive work) at the expense of hardware ($\zeta N + K$, gradually dominated by K). Hence, in the second stage of development of an industrial economy, human cognitive work becomes increasingly scarce and thus increasingly well remunerated, raising the returns to education and the skill premium, and setting up a secular upward trend in wages.[19] In the hypothetical limit of $A \to \infty$, $K \to \infty$ and $h \to \bar{h}$ without a digital revolution, the economy tends to a balanced growth path, along which $Y = F(K, A\bar{h}N)$, the hardware (=capital) share stabilizes around some intermediate value $\bar{\pi}_X \in (0, 1)$, and the economy respects Kaldor's facts (Kaldor, 1961).

Stage 3. Digital production. The first stage of development of a digital economy features gradual *automation* of production: accumulation of pre-programmed software Ψ gradually reduces the role of human cognitive work H in software. Given the substitutability of these two factors, the cognitive labor share goes down whereas the pre-programmed software share goes up. (And if data and software rents are not separately accounted, also firms' profit shares and markups go up, as documented e.g. by Barkai (2017); De Loecker and Eeckhout (2018).) This is the world of today, where disruptive digital technologies fuel the "rise of the global 1%".

The hardware-software model predicts a change of this secular trend in the future, though. It expects that due to exponential technological progress in A, systematic improvements in AI skill ψ, and progressing automation, hardware will gradually become the bottleneck of global development, a key factor constraining the pace of further economic growth. Consequently the revenues will be increasingly directed towards (programmable) hardware, and the software share π_S will set on a secular downward trend. In the hypothetical limit of $K \to \infty, \chi \to \bar{\chi}, \psi \to \bar{\psi}$, assuming the absence of a next technological revolution, $Y = K \cdot F(1, A\bar{\psi}\bar{\chi})$ and the hardware share will tend to unity. At that point in time, though, only a negligible fraction of the remuneration will be going to human workers because all human skills will by then have been fully mechanized and automated.

Hypothetical stage 4. Post-digital production. Perhaps the functional distribution of income becomes less of an issue in a world where neither hardware nor software requires any human input, but nevertheless one may observe that the episode of "new mechanization" (replacement of K with M in hardware) would incur a dynamic that is largely similar to the one following the Industrial Revolution. Namely,

[19] As Galor and Moav (2006) put it, "The accumulation of physical capital in the early stages of industrialization enhanced the importance of human capital in the production process and generated an incentive for the capitalists to support the provision of public education for the masses, planting the seeds for the demise of the existing class structure".

accumulation of M would systematically decrease the role of K in hardware, so that the share of K would go down whereas the share of M would go up. Next, if all software would be able to scale with M then its share would remain low; if not then it would become increasingly scarce and its share of output would go up.

4 The R&D Equation

Technological progress due to purposeful R&D activities is widely acknowledged as a key driver of long-run growth in output per worker in the industrial and early digital era. Due to the non-rivalrous character of technological ideas, they act a source of increasing returns to scale (Romer, 1986, 1990), allowing output to grow even when the use of inputs is constant over time. The exact specification of the R&D process at the macroeconomic level is however subject to dispute. In particular, and perhaps somewhat surprisingly, most of the existing R&D-based growth literature assumes that researchers' labor is the only input in the R&D process (Romer, 1990; Jones, 1995, 1999; Ha and Howitt, 2007). Alternatively, a few studies embrace the "lab equipment" specification of the R&D process, conditioning R&D output on overall R&D spending (Rivera-Batiz and Romer, 1991; Bloom, Jones, Van Reenen, and Webb, 2017; Kruse-Andersen, 2017). In reality, however, productivity of the R&D sector depends not just on the labor of researchers but increasingly also on the services of *R&D capital*. Modern R&D capital may range from modest offices at university campuses or computers at researchers' laps to such exquisite machinery as the Very Large Telescope (VLT) and the Large Hadron Collider (LHC). Historically, the practicality and complexity of research equipment has undergone systematic, cumulative changes over the centuries. The difference in usefulness of Ptolemy's astrolabe, Galileo's telescope and the VLT is breathtaking, and so is to think how early statisticians could compute correlations and run regressions without relying on computers.

4.1 Setup

Consistently with the hardware–software model, I postulate that R&D output should be a function of two inputs to the R&D process: hardware X and software S. Hardware includes R&D capital alongside human physical labor. Software encompasses all the sophisticated and ingenuous ideas supplied by scientists and technical personnel, as well as – increasingly – code encapsulated in pre-programmed software.

Intuitively, the difference between the production process and the R&D process is that the latter tends to involve relatively less physical action and more sophisticated instructions. R&D is also not bound by the thermodynamical requirement that an essential fraction of its output must be material; in fact probably most if not all of its output comes in the form of information. Yet, from the conceptual perspective hardware must be considered essential also in R&D. After all, even pure thinking is in fact information processing carried out in the thinker's brain – so it needs some hardware, too; and the further we go from genuinely abstract, philosophical reflection towards more applied R&D, the more actual physical action is necessary, such as laboratory experiments, survey data collection, model building, or prototype testing.

The hardware–software framework implicitly assumes that there is no qualitative difference between human thought and computer software in digital-era R&D processes. In line with Brynjolfsson and McAfee (2014) I hypothesize that *ideation*, creativity and intuition represent sophisticated pattern recognition. Thus there is no theoretical argument precluding the possibility that R&D will also be subject to gradual automation in the digital era. Today AI is already used in e.g., genome sequencing, not to mention web browser engines, which are of enormous help to modern researchers. In the future, AI may revolutionize R&D by not just helping people in answering research questions, but also in asking new ones.

Formally I postulate that the idea production function should also obey the general equation (2):[20]

$$\dot{A} = \Phi(X, S) = \Phi(\zeta N + K, A(hN + \psi\chi K)). \tag{12}$$

It is assumed that Φ is increasing and concave in both factors, X and S. The characterization of returns to scale is uncertain, however, as there may be important spillover effects and duplication externalities in R&D, the magnitude of which is subject to dispute (Jones, 1999; Ha and Howitt, 2007; Madsen, 2008; Kruse-Andersen, 2017; Bloom, Jones, Van Reenen, and Webb, 2017).

4.2 R&D Across Stages of Economic Development

Let me now discuss how the overarching hardware–software framework specializes to deal with the realities of consecutive eras of economic development.

Stage 1. Pre-industrial R&D. In a pre-industrial economy, R&D was carried out mainly by individual scholars and their disciples. R&D output was generated essentially from their thought and simple experiments, with little or no aid of R&D capital. Setting $K = 0$ and $\chi = 0$ in (12) yields:

$$\dot{A} = \Phi(X, S) = \Phi(\zeta N, AhN). \tag{13}$$

[20] In order to better describe the early millennia of human history, equation (12) could be augmented with knowledge depreciation. As the focus here is on the more recent centuries, after the development of writing and the printing press, which massively reduced depreciation of aggregate human knowledge, I set this complication aside.

Hence, under gross complementarity of hardware and software, the pool of technological opportunity was gradually depleted and "ideas were getting harder to find" (Olsson, 2005; Bloom, Jones, Van Reenen, and Webb, 2017). The model implies that in the absence of R&D capital, in the counterfactual scenario of $A \to \infty$ with a fixed N the knowledge increment $\dot{A}$ would tend to a positive constant and the rate of technological progress $\dot{A}/A$ – to zero.

Stage 2. R&D in the industrial era. In an industrial economy, R&D output was produced increasingly in universities, laboratories, specialized research units and corporate R&D divisions. More and more specialized machines were employed in order to advance the state of knowledge. All physical actions were, however, dependent on the instructions provided by scientists and technicians: there was no programmable hardware and no pre-programmed software yet ($\chi = 0$ and thus $\Psi = 0$). Transforming (12), the following form is obtained:

$$\dot{A} = \Phi(X, S) = \Phi(\zeta N + K, AhN). \tag{14}$$

In the hypothetical limit of full mechanization and skill satiation, $K \to \infty$ and $h \to \bar{h}$, the model implies that $\dot{A} = \Phi(K, \bar{h}AN)$, where $\bar{h}$ is the upper limit for human skills. Thus $\dot{A}/A$ is a decreasing function of A, and again "ideas are getting harder to find".

Moreover, under the additional assumption that Φ exhibits constant returns to scale and capital is accumulated in the standard way (Solow, 1956), we can also derive a clear-cut prediction about the results of interplay between technological progress and (R&D and non-R&D) capital accumulation in the long run. In such a case the economy would approach an asymptotic balanced growth path where K and A grow at the same rate:

$$g_A = \frac{\dot{A}}{A} = \Phi\left(\frac{K}{A}, \bar{h}N\right), \tag{15}$$

$$g_K = \frac{\dot{K}}{K} = \bar{s}\frac{Y}{K} - \delta = \bar{s}F\left(1, \frac{A}{K}\bar{h}N\right) - \delta, \tag{16}$$

where $\bar{s} \in (0, 1]$ is the long-run limit of the savings rate and $\delta > 0$ represents the rate of capital depreciation. Hence, in the counterfactual scenario of asymptotically balanced growth without a digital revolution, any potential increases in R&D employment would tend to increase the pace of technological progress only up to a point, after which that rate would be pinned by the scarce factor, K/A.

In the absence of pre-programmed software in the economy, R&D is the key source of economic growth, whereas accumulation of R&D capital is the key mechanism allowing to sustain it.

Stage 3. R&D in the digital era. In the early days of the digital era in which we are living today, human research skills are increasingly augmented with sophisticated

R&D hardware and some of the more tedious research tasks are gradually automated. This process may accelerate fast in the future after AI algorithms become sufficiently advanced to meaningfully contribute also to non-routine research tasks.

In the digital era, equation (12) holds in its general form. The hypothetical limit of full automation implies:

$$\dot{A} = \Phi(K, A\bar{\psi}\bar{\chi}K). \tag{17}$$

If additionally Φ exhibits constant returns to scale then the economy tends to an asymptotic balanced growth path where K and A grow at the same rate:

$$g_A = \frac{\dot{A}}{A} = \Phi\left(\frac{K}{A}, \bar{\psi}\bar{\chi}K\right), \tag{18}$$

$$g_K = \frac{\dot{K}}{K} = \bar{s}\frac{Y}{K} - \delta = \bar{s}F\left(1, A\bar{\psi}\bar{\chi}\right) - \delta. \tag{19}$$

Hence, in the hypothetical long-run limit the accumulation of programmable hardware χK would tend to increase the pace of technological progress only up to a point, after which it would be pinned by the scarce factor, K/A.

The hardware–software model predicts that after the global economy has passed through a digital revolution, accumulation of programmable hardware must eventually become the unique engine of long-run growth. In a world where software is no longer pinned to the size of the human population and instead is able to scale with hardware, technological progress will eventually cease to be the key contributor to growth.

5 Production, R&D and Growth in the Digital Era: A CES Example

Let me now provide a more detailed treatment of the impact of the Digital Revolution on global production, R&D and growth under the proposed framework. To this end, I will specifically assume that both production functions, F and Φ, take the normalized CES form (Klump, McAdam, and Willman, 2012), whereas capital accumulation follows the standard equation of motion due to Solow (1956). What follows is a two-sector growth model with two interlinked growth engines, capital accumulation and R&D. Neither of them is able to drive long-run growth alone: capital accumulation is not sufficient because of decreasing returns under any fixed level of technology; and R&D is not sufficient because its operations require R&D capital (unlike endogenous growth models such as Romer (1990) or Jones (1995) where human cognitive work was the only essential factor in R&D).

The model consists of the following equations:

$$X = \zeta N + K, \tag{20}$$

$$S = A(hN + \psi\chi K), \tag{21}$$

$$Y = Y_0 \left(\pi_0 \left(\frac{u_X X}{u_{X0} X_0} \right)^{\xi} + (1-\pi_0) \left(\frac{u_S S}{u_{S0} S_0} \right)^{\xi} \right)^{\frac{1}{\xi}}, \tag{22}$$

$$\dot{A} = \dot{A}_0 \left(\gamma_0 \left(\frac{(1-u_X)X}{(1-u_{X0})X_0} \right)^{\mu} + (1-\gamma_0) \left(\frac{(1-u_S)S}{(1-u_{S0})S_0} \right)^{\mu} \right)^{\frac{1}{\mu}}, \tag{23}$$

$$\dot{K} = sY - \delta K, \tag{24}$$

where $s \in [0,1]$ is the savings rate, $u_X, u_S \in [0,1]$ are the shares of hardware and software, respectively, allocated to the production sector, and $(1-u_X), (1-u_S)$ are the respective shares allocated to R&D. The parameter $\xi < 0$ captures the degree of substitutability between hardware and software in production, and $\mu < 0$ – in R&D. The parameters with subscript 0 are normalization constants. Population $N > 0$ is assumed constant, which is a realistic assumption for the long run given that United Nations population projections suggest that global population will plateau within the next century.

To concentrate uniquely on technological underpinnings of long-run economic growth and not the role of preferences or institutions, I consider a rule-of-thumb allocation of resources à la Solow (1956); Mankiw, Romer, and Weil (1992) where the savings rate s and the shares u_X, u_S are exogenous and constant. Allowing them to be set optimally by utility-maximizing decision makers is left for another research.

This framework allows me to provide a comparison of two polar scenarios: (i) without any digital revolution ($\chi = 0$), and (ii) with a digital revolution, eventually leading to full automation.

Industrial-era economy without automation. In an economy without programmable hardware and pre-programmed software, as the stock of capital tends to infinity and as $h \to \bar{h}$ (growth in the average level of education flattens out), one may approximate $X \approx K$ (full mechanization) and $S \approx A\bar{h}N$. Inserting these approximations into the system (20)–(24) and setting a constant population size $N = N_0$ yields the following system of equations describing the balanced growth path of the economy:

$$\frac{Y}{A} = \frac{Y_0}{A_0} \left(\pi_0 \left(\frac{u_X}{u_{X0}} \frac{K}{A} \frac{A_0}{K_0} \right)^{\xi} + (1-\pi_0) \left(\frac{u_S}{u_{S0}} \frac{\bar{h}}{h_0} \right)^{\xi} \right)^{\frac{1}{\xi}}, \tag{25}$$

$$g = g_0 \left(\gamma_0 \left(\frac{(1-u_X)}{(1-u_{X0})} \frac{K}{A} \frac{A_0}{K_0} \right)^{\mu} + (1-\gamma_0) \left(\frac{(1-u_S)\bar{h}}{(1-u_{S0})h_0} \right)^{\mu} \right)^{\frac{1}{\mu}}, \tag{26}$$

$$g = s\frac{Y}{K} - \delta, \tag{27}$$

$$\frac{Y}{A} = \frac{Y}{K}\frac{K}{A}. \tag{28}$$

This is a four-equation system in four stationary variables: the growth rate g and three ratios, Y/A, Y/K and K/A. Additional calculus uncovers that the long-run economic growth rate g depends on the key parameters of the model, s, u_X and u_S. The dependence of g on s is unambiguously positive, whereas growth effects of the latter two parameters are ambiguous.

Along the balanced growth path of the industrial economy without automation, the economy respects Kaldor (1961) facts: it grows at a steady rate g while the "great ratios" $(K/Y, C/Y)$ and factor shares are constant.

Digital-era economy with full automation. In an economy with programmable hardware and pre-programmed software, as the stock of capital tends to infinity, and $\chi \to \bar{\chi}, \psi \to \bar{\psi}$, one may approximate $X \approx K$ and $S \approx A\bar{\psi}\bar{\chi}K$. This underscores that in the limit of full mechanization and automation, production and R&D become entirely decoupled from the employed human population. Inserting these approximations into the system (20)–(24) and letting $A \to \infty$ yields the following asymptotic balanced growth path of this economy:

$$g = \frac{\dot{Y}}{Y} = \frac{\dot{K}}{K} = \frac{\dot{A}}{A} = s\pi_0^{\frac{1}{\xi}} \left(\frac{u_X}{u_{X0}} \frac{Y_0}{K_0} \right) - \delta, \tag{29}$$

$$\frac{Y}{K} = \pi_0^{\frac{1}{\xi}} \frac{u_X}{u_{X0}} \frac{Y_0}{K_0}, \tag{30}$$

$$\frac{K}{A} = \gamma_0^{-\frac{1}{\mu}} \left(\frac{1-u_{X0}}{1-u_X} \right) \frac{g}{g_0}. \tag{31}$$

Hence, this scenario leads to an AK-type model of fully endogenous growth (Jones and Manuelli, 1990; Barro and Sala-i-Martin, 2003). The accumulation of programmable hardware becomes the unique engine of growth because it does double duty once software is able to scale up to hardware. The impact of R&D on growth eventually vanishes. The parameters positively affecting the long-run growth rate are (i) the savings rate s, and (ii) the share of hardware in production u_X. The allocation of software u_S becomes irrelevant in the limit because if software is able to scale with hardware, it is ultimately only the hardware that determines the pace of economic growth. Accordingly, as the impact of R&D on growth gradually disappears, in the limit it does not make sense to allocate any more hardware to R&D.

Along the asymptotic balanced growth path of the digital economy with full automation, the economy respects the Kaldor (1961) facts of constancy of the growth rate g and the "great ratios" $(K/Y, C/Y)$, but the software share falls to zero.

6 Discussion

6.1 Key Concepts and Misconceptions of the Digital Era

In the current paper I have carried out some basic conceptual work needed by economic growth theory to achieve progress in modeling the realities of the digital era. The key contribution of the proposed hardware–software model is to formalize production processes observed across the human history, with a specific focus on the effects of the Digital Revolution. In particular the current paper provides a conceptually consistent approach to delineating such key concepts as mechanization, automation and the adoption of ICT and AI.

Viewed through the lens of the hardware–software model:

- *Mechanization* of production consists in replacing human (and animal) physical labor with machines (K in place of L) within hardware. Large-scale mechanization is observed since the Industrial Revolution (≈1800 CE onwards). Mechanization applies to physical actions but not the instructions defining them.

- *Automation* of production consists in replacing humans with pre-programmed software in providing instructions to machines (Ψ in place of H), i.e., within software. Automation is observed since the Digital Revolution (≈1980 CE onwards) when information technologies first came into use as general purpose technologies (Bresnahan and Trajtenberg, 1995). Automation pertains to cases where a task, previously involving human thought and decisions, is carried out entirely by machines without any human intervention. Routine tasks (both physical and cognitive) are typically among the first to be automated (Autor and Dorn, 2013).

Historically mechanization preceded automation. Therefore the automation processes of the digital era frequently affect tasks where no human physical labor is needed anymore. This ordering is however not obligatory. A fun example of automation without mechanization is when you walk around town blindly following the instructions of your GPS.

The hardware–software model is also helpful in providing an economic frame for the concepts of ICT and AI.

- *Information and communication technology* (ICT) is a special type of physical capital that has the ability to store and run code. ICTs constitute a breakthrough compared to non-programmable machines because they allow to replace humans in providing instructions. Code, once programmed, can be run multiple times, also concurrently on many machines, without the need of any human intervention. Hence ICTs, which can be roughly equated with programmable hardware, were necessary for initiating automation.

- *Artificial intelligence (AI)* is a special type of pre-programmed software that has the ability to learn from data. In contrast to "traditional" software which consists of a fixed set of instructions (e.g., if–then loops), artificial intelligence can improve its performance based on experience and new information. This happens even under a static architecture of AI algorithms, though it is conceivable that such algorithms may also modify their own architecture while heading towards self-improvement. The advantage of machine learning over human learning is that networked pieces of equipment can effectively pool their data whereas humans cannot. The development of AI opens new opportunities for speeding up automation because AI allows to substitute humans in non-routine tasks as well (Brynjolfsson, Rock, and Syverson, 2019). According to Agrawal, Gans, and Goldfarb (2017), while computers drastically lowered the costs of computing (arithmetic), AI drastically lowers the costs of *prediction*.

In light of the above discussion, it is a misconception to identify computers and robots with their hardware. To be useful in generating value added, computers, robots, smartphones and other ICTs must also be provided with appropriate instructions, stemming either from human cognitive work or pre-programmed software.

Another frequent misconception is to automatically associate AI with robots. AI is software that can learn from data. This software may indeed provide instructions to robots, but also to conventional computers, smartphones and other programmable devices.

It is also rather problematic to identify AI development with automation, because automation may proceed also without AI, as it has been the case for decades e.g. in the auto industry, and AI – especially at initial stages of development – may be complementary to some human skills such as judgment (Agrawal, Gans, and Goldfarb, 2017).

Last but not least, automation also should not be conflated with mechanization. This is done, for example, in the famous question "will humans go the way of horses?" (Brynjolfsson and McAfee, 2014), that is whether human work will be eventually fully replaced by machines. The answer is: as far as physical labor is concerned, we have long gone the way of horses; for cognitive tasks (for which horses are of no use) this has not been the case, at least not yet. By the same token, it is false comfort to say that the history of the Industrial Revolution teaches us that when jobs are destroyed, new ones are bound to emerge. It only teaches us that when physical labor is mechanized, additional workers will be demanded in cognitive occupations, but it tells us nothing about cognitive occupations being automated.

Mechanization and automation are also habitually conflated when using terms like the "Fourth Industrial Revolution" or "Industry 4.0" (Schwab, 2016).

6.2 Software Capabilities and the Future of the World Economy

According to the hardware–software model the development of sophisticated pre-programmed software such as AI has a decisive impact on the long-run growth path of the world economy. This result critically depends, however, on the assumption that there is no upper bound for the accumulation of pre-programmed software Ψ relative to human cognitive work H. But is this a reasonable assumption? This is a dilemma that has both an extensive and an intensive margin. At the extensive margin, the question is whether the pace of growth in aggregate computing power, data storage and bandwidth (χK) can remain systematically higher than the rate of accumulation of human capital (hN). At the intensive margin, in turn, the question pertains to the domain of the AI skill level ψ: can AI potentially replace people in all tasks, including R&D, inventing new tasks and designing AI? Can AI potentially achieve superhuman performance across a broad array of tasks and gain sufficient adaptability and versatility to be able to endogenously expand the breadth of its expertise?

If both answers are "yes", there will be no clear upper bound for automation. If both are "no", at some point automation will surely stop. If only the first one is answered affirmatively, though, the possibility of full automation depends on whether people will forever maintain an edge over AI in at least some essential tasks. Depending on the answers to these questions the hardware–software model places the future of the world economy on a spectrum between secular stagnation and a technological singularity.[21]

- *Secular stagnation.* If both answers are negative, so that there is a firm upper bound for automation, and moreover R&D (the function Φ) is characterized by decreasing returns to scale, then economic growth will gradually slow down and eventually the world economy will settle in a steady state or grow sub-exponentially.

- *Balanced growth with bounded automation.* If both answers are negative but R&D (the function Φ) is characterized by constant returns to scale, there will still be a firm upper bound for automation. In such a scenario, however, the economic growth rate will converge to a constant and eventually the world

[21] The scenarios below are formulated under "technological determinism", i.e., assuming that all which is technologically feasible will surely be attained. However, there may be specific preferences or institutions which would preclude full automation, limit technological progress, etc.

economy will reach a balanced growth path, along which further growth will be driven by technological progress and sustained by the accumulation of (R&D and non-R&D) capital. The long-run growth rate will be pinned to the growth in aggregate human cognitive work AhN (perhaps in the order of 2–3% per annum).

- *Race against the machine.* Qualitatively the same results as above are obtained also in the case where the first answer is positive, the second answer is negative, but where people will always keep an edge over AI in some essential tasks (such as R&D, inventing new tasks or building AI, Acemoglu and Restrepo, 2018).
- *Balanced growth with unbounded automation.* If the first answer is positive, the second answer is negative, and people will eventually lose their edge over AI in all essential tasks, then there will be no upper bound for automation. The economic growth rate will then eventually converge to a constant and in the absence of further technological revolutions the world economy will reach a balanced growth path, along which further growth will be driven solely by the accumulation of programmable hardware. The long-run growth rate will then be no longer pinned to growth in aggregate human cognitive work, and thus will be visibly faster (perhaps in the order of 20–30% per annum or more).
- *Technological singularity.* Qualitatively the same results as above are obtained also in the case where both answers are positive. In such a scenario, though, in finite time the world will reach *technological singularity*, or "AI takeover". From that moment onwards, AI will exhibit superhuman cognitive performance in all essential tasks, and consequently will take over all important decisions related to the functioning of the world economy (Kurzweil, 2005; Nordhaus, 2017; Aghion, Jones, and Jones, 2019).[22]

6.3 Technological Singularity?

So is technological singularity feasible? Will people one day lose control over the critical decisions in the world economy? The answer depends crucially on two issues. First, is *ideation* a sophisticated incarnation of pattern recognition or a qualitatively different feature? Can AI be creative, imaginative and insightful in the way humans can be? Preliminary evidence suggests that the answer is likely affirmative. Even some of the contemporary AI algorithms can indeed be perceived as creative, e.g., in devising innovative strategies in chess and Go (DeepMind's AlphaZero, Silver, Hubert, Schrittwieser, et al., 2018), drawing artistic pictures (Schmidhuber, 2009a),

[22]Consequences of technological singularity extend way beyond the economy. Such an event will surely have tremendous psychological, political and even *existential* implications for the humankind (see e.g. Hanson and Yudkowsky, 2013; Bostrom, 2014; Harari, 2017).

or composing music (Amper Music, IBM's Watson Beat, Google's Magenta, AIVA). Arguments have also been formulated that the lines between creativity, insight and complexity are actually rather arbitrary and subjective (Dennett, 2017; Tegmark, 2017).

Second, how high are the *returns to cognitive reinvestment* in AI? (Yudkowsky, 2013) How efficient will the future AI be in re-designing itself and its environment in order to improve its cognitive capacity? Humans are in this regard limited by our inability to rewire our brains, and so we circumvent this limitation by increasingly relying on external memory, data collection equipment, and computational power. We also increasingly pool our resources by working in ever larger teams whose members have increasingly specialized sets of skills. As our knowledge set is growing but our brains are not, interdisciplinary "Renaissance Men" are long gone (Jones, 2009). Unfortunately, speed and accuracy of our interpersonal communication are far from perfect, and thus we may be missing plenty of interdisciplinary insights. AI algorithms running on fast computers, in contrast, communicate extremely fast and without error. They also by far surpass us in terms of speed and serial depth of computation (Hanson and Yudkowsky, 2013). In contrast to humans, AI is also (so far, theoretically) able to recursively rewrite its code provided that it is able to prove that the rewrite is beneficial (Schmidhuber, 2009b). Hence, although there are no hard data yet which would allow to quantify the returns to cognitive reinvestment in AI, preliminary evidence suggests potentially high overall AI capabilities and motivates the baseline parametrization used in the current paper.

The main disadvantage of modern-day AI algorithms, though, is that they are markedly lagging behind the human brain in terms of versatility and adaptivity. If this issue is resolved, we will observe a rapid buildup of AI skills, and perhaps even an uncontrolled intelligence explosion (Hanson and Yudkowsky, 2013; Bostrom, 2014). Consistently with the hardware–software model, the world will then be facing technological singularity.[23]

* * *

Future work on the hardware–software model should forge a link between the proposed conceptual framework and general-equilibrium modeling of economic growth. It is important to identify the equilibrium forces determining the extent of automation and to quantify the timing at which AI development becomes critical for further economic growth. One could also review alternative scenarios, such as the one where R&D could be carried out without R&D capital or where AI software does not scale

[23]By contrast, singularity understood as a vertical asymptote in the level of GDP, i.e., arbitrarily high production in finite time, is not possible. Given that a non-degenerate fraction of output must be material to sustain the hardware, such a scenario would be inconsistent with the laws of thermodynamics.

proportionally to hardware. Another promising line of work would be to analyze complex tasks within the hardware–software model in order to quantify the extent to which human cognitive work and AI can be complementary on the run-up to full automation.

References

Acemoglu, D. (2003): "Labor-and Capital-Augmenting Technical Change," *Journal of the European Economic Association*, 1, 1–37.

——— (2009): *Introduction to Modern Economic Growth.* Princeton University Press.

Acemoglu, D., and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4, chap. 12, pp. 1043–1171. Elsevier.

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

Aghion, P., and P. Howitt (1992): "A Model of Growth Through Creative Destruction," *Econometrica*, 60, 323–351.

Aghion, P., B. F. Jones, and C. I. Jones (2019): "Artificial Intelligence and Economic Growth," in *The Economics of Artificial Intelligence: An Agenda*, ed. by A. Agrawal, J. Gans, and A. Goldfarb, pp. 237–282. University of Chicago Press.

Agrawal, A., J. S. Gans, and A. Goldfarb (2017): "What to Expect From Artificial Intelligence," *MIT Sloan Management Review*, Spring 2017.

Andrews, D., C. Criscuolo, and P. N. Gal (2016): "The Global Productivity Slowdown, Technology Divergence and Public Policy: A Firm Level Perspective," Working party no. 1 on macroeconomic and structural policy analysis, OECD.

Arntz, M., T. Gregory, and U. Zierahn (2016): "The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis," OECD Social, Employment and Migration Working Paper No 189, OECD Publishing, Paris.

Arpaia, A., E. Pérez, and K. Pichelmann (2009): "Understanding Labour Income Share Dynamics in Europe," Working paper, European Commission, European Economy Paper No. 379.

Arute, F., K. Arya, R. Babbush, et al. (2019): "Quantum Supremacy Using a Programmable Superconducting Processor," *Nature*, 574, 505–510.

Aum, S., S. Y. Lee, and Y. Shin (2018): "Computerizing Industries and Routinizing Jobs: Explaining Trends in Aggregate Productivity," *Journal of Monetary Economics*, 97, 1–21.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2017): "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper No. 23396, NBER.

Autor, D. H., and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5), 1553–97.

Barkai, S. (2017): "Declining Labor and Capital Shares," Job market paper, University of Chicago.

Barro, R. J., and X. X. Sala-i-Martin (2003): *Economic Growth*. MIT Press.

Benzell, S. G., L. J. Kotlikoff, G. LaGarda, and J. D. Sachs (2015): "Robots Are Us: Some Economics of Human Replacement," Working Paper No. 20941, NBER.

Berg, A., E. F. Buffie, and L.-F. Zanna (2018): "Should We Fear the Robot Revolution? (The Correct Answer is Yes)," *Journal of Monetary Economics*, 97, 117–148.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb (2017): "Are Ideas Getting Harder to Find?," Working paper 23782, National Bureau of Economic Research.

Bostrom, N. (2014): *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

Bresnahan, T. F., and M. Trajtenberg (1995): "General Purpose Technologies: Engines of Growth?," *Journal of Econometrics*, 65, 83–108.

Brynjolfsson, E., and A. McAfee (2014): *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W.W. Norton & Co.

Brynjolfsson, E., D. Rock, and C. Syverson (2019): "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," in *The Economics of Artificial Intelligence: An Agenda*, ed. by A. Agrawal, J. Gans, and A. Goldfarb, pp. 23–57. University of Chicago Press.

Caselli, F., and W. J. Coleman (2006): "The World Technology Frontier," *American Economic Review*, 96, 499–522.

De Loecker, J., and J. Eeckhout (2017): "The Rise of Market Power and the Macroeconomic Implications," Working Paper 23687, National Bureau of Economic Research.

——— (2018): "Global Market Power," Working Paper 24768, National Bureau of Economic Research.

DeCanio, S. J. (2016): "Robots and Humans – Complements or Substitutes?," *Journal of Macroeconomics*, 49, 280–291.

Dennett, D. (2017): *From Bacteria to Bach and Back: The Evolution of Minds.* W. W. Norton & Company.

Diez, F., D. Leigh, and S. Tambunlertchai (2018): "Global Market Power and Its Macroeconomic Implications," Working Paper No. 18/137, International Monetary Fund.

Elsby, M., B. Hobijn, and A. Sahin (2013): "The Decline of the U.S. Labor Share," *Brookings Papers on Economic Activity*, 47(2), 1–63.

Frey, C. B., and M. Osborne (2017): "The Future of Employment: How Susceptible Are Jobs to Computerisation?," *Technological Forecasting and Social Change*, 114, 254–280.

Galor, O. (2005): "From Stagnation to Growth: Unified Growth Theory," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, pp. 171–293. North-Holland.

——— (2011): *Unified Growth Theory.* Princeton: Princeton University Press.

Galor, O., and O. Moav (2006): "Das Human-Kapital: A Theory of the Demise of the Class Structure," *Review of Economic Studies*, 73, 85–117.

Galor, O., and D. N. Weil (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 90, 806–828.

Georgescu-Roegen, N. (1971): *The Entropy Law and the Economic Process.* Harvard University Press.

——— (1975): "Energy and Economic Myths," *Southern Economic Journal*, 41, 347–381.

Gillings, M. R., M. Hilbert, and D. J. Kemp (2016): "Information in the Biosphere: Biological and Digital Worlds," *Trends in Ecology and Evolution*, 31, 180–189.

Gordon, R. J. (1990): *The Measurement of Durable Goods Prices*. University of Chicago Press.

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

Grace, K. (2013): "Algorithmic Progress in Six Domains," Technical report 2013-3, Berkeley, CA: Machine Intelligence Research Institute.

Graetz, G., and G. Michaels (2018): "Robots at Work," *Review of Economics and Statistics*, 100, 753–768.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997): "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 87, 342–362.

Growiec, J. (2012): "The World Technology Frontier: What Can We Learn from the US States?," *Oxford Bulletin of Economics and Statistics*, 74, 777–807.

——— (2013): "A Microfoundation for Normalized CES Production Functions with Factor-Augmenting Technical Change," *Journal of Economic Dynamics and Control*, 37, 2336–2350.

Growiec, J. (2018): "Factor-Specific Technology Choice," *Journal of Mathematical Economics*, 77, 1–14.

Growiec, J., and J. Mućk (2018): "Isoelastic Elasticity of Substitution Production Functions," *Macroeconomic Dynamics*, p. forthcoming.

Ha, J., and P. Howitt (2007): "Accounting for Trends in Productivity and R&D: A Schumpeterian Critique of Semi-Endogenous Growth Theory," *Journal of Money, Credit and Banking*, 39(4), 733–774.

Hanson, R. (2000): "Long-Term Growth as a Sequence of Exponential Modes," Working paper, George Mason University.

Hanson, R., and E. Yudkowsky (2013): *The Hanson-Yudkowsky AI-Foom Debate*. Machine Intelligence Research Institute.

Harari, Y. N. (2017): *Homo Deus: A Brief History of Tomorrow*. Vintage.

Hemous, D., and M. Olsen (2018): "The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality," Working paper, University of Zurich.

Henderson, D. J., and R. R. Russell (2005): "Human Capital and Convergence: A Production–Frontier Approach," *International Economic Review*, 46, 1167–1205.

Hercowitz, Z. (1998): "The 'Embodiment' Controversy: A Review Essay," *Journal of Monetary Economics*, 41, 217–224.

Hilbert, M., and P. López (2011): "The World's Technological Capacity to Store, Communicate, and Compute Information," *Science*, 332, 60–65.

Jones, B. F. (2009): "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? ," *Review of Economic Studies*, 76, 283–317.

Jones, C. I. (1995): "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103, 759–84.

——— (1999): "Growth: With or Without Scale Effects?," *American Economic Review*, 89(2), 139–144.

——— (2002): "Sources of U.S. Economic Growth in a World of Ideas," *American Economic Review*, 92, 220–239.

——— (2005a): "Growth and Ideas," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. North-Holland.

——— (2005b): "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics*, 120, 517–549.

Jones, C. I., and J. Kim (2018): "A Schumpeterian Model of Top Income Inequality," *Journal of Political Economy*, 126, 1785–1826.

Jones, C. I., and P. M. Romer (2010): "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital," *American Economic Journal: Macroeconomics*, 2, 224–245.

Jones, L. E., and R. E. Manuelli (1990): "A Convex Model of Equilibrium Growth: Theory and Policy Implications," *Journal of Political Economy*, 98, 1008–1038.

Jorgenson, D. W. (1995): *Productivity. Volume 1: Postwar U.S. Economic Growth*. MIT Press.

——— (2005): "Accounting for Growth in the Information Age," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. North-Holland.

Kaldor, N. (1961): "Capital Accumulation and Economic Growth," in *The Theory of Capital*, ed. by L. A. Lutz, and D. C. Hague, pp. 177–222. Palgrave Macmillan.

Karabarbounis, L., and B. Neiman (2014): "The Global Decline of the Labor Share," *Quarterly Journal of Economics*, 129(1), 61–103.

Klein, L. R., and R. F. Kosobud (1961): "Some Econometrics of Growth: Great Ratios of Economics," *Quarterly Journal of Economics*, 75, 173–198.

Klump, R., P. McAdam, and A. Willman (2007): "Factor Substitution and Factor Augmenting Technical Progress in the US," *Review of Economics and Statistics*, 89, 183–192.

——— (2012): "Normalization in CES Production Functions: Theory and Empirics," *Journal of Economic Surveys*, 26, 769–799.

Koop, G., J. Osiewalski, and M. F. J. Steel (1999): "The Components of Output Growth: A Stochastic Frontier Analysis," *Oxford Bulletin of Economics and Statistics*, 61, 455–487.

——— (2000): "Measuring the Sources of Output Growth in a Panel of Countries," *Journal of Business and Economic Statistics*, 18, 284–299.

Kruse-Andersen, P. K. (2017): "Testing R&D-Based Endogenous Growth Models," Working paper, University of Copenhagen.

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, 1029–1054.

Kumar, S., and R. R. Russell (2002): "Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence," *American Economic Review*, 92, 527–548.

Kurzweil, R. (2005): *The Singularity is Near*. New York: Penguin.

León-Ledesma, M. A., P. McAdam, and A. Willman (2010): "Identifying the Elasticity of Substitution with Biased Technical Change," *American Economic Review*, 100, 1330–1357.

Madsen, J. (2008): "Semi-endogenous Versus Schumpeterian Growth Models: Testing the Knowledge Production Function Using International Data," *Journal of Economic Growth*, 13, 1–26.

Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407–437.

McAdam, P., and A. Willman (2018): "Unraveling the Skill Premium," *Macroeconomic Dynamics*, 22, 33–62.

Milanović, B. (2016): *Global Inequality: A New Approach for the Age of Globalization*. Harvard University Press.

Mućk, J. (2017): "Elasticity of Substitution Between Labor and Capital: Robust Evidence from Developed Economies," Working paper no. 271, Narodowy Bank Polski.

Nordhaus, W. D. (2017): "Are We Approaching an Economic Singularity? Information Technology and the Future of Economic Growth," Working paper, Cowles Foundation, Yale University.

Olsson, O. (2005): "Technological Opportunity and Growth," *Journal of Economic Growth*, 10, 31–53.

Piketty, T. (2014): *Capital in the Twenty-First Century*. Harvard University Press, Translated from French by Arthur Goldhammer.

Piketty, T., and G. Zucman (2014): "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010," *Quarterly Journal of Economics*, 129(3), 1255–1310.

Rivera-Batiz, L. A., and P. M. Romer (1991): "Economic Integration and Endogenous Growth," *The Quarterly Journal of Economics*, 106(2), 531–555.

Romer, P. M. (1986): "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94(5), 1002–1037.

——— (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, S71–S102.

Sachs, J. D., S. G. Benzell, and G. LaGarda (2015): "Robots: Curse or Blessing? A Basic Framework," Working Paper Np. 21091, NBER.

Schmidhuber, J. (2009a): "Driven by Compression Progress: A Simple Principle Explains Essential Aspects of Subjective Beauty, Novelty, Surprise, Interestingness, Attention, Curiosity, Creativity, Art, Science, Music, Jokes," in *Anticipatory Behavior in Adaptive Learning Systems, from Sensorimotor to Higher-level Cognitive Capabilities*, ed. by G. Pezzulo, M. V. Butz, O. Sigaud, and G. Baldassarre. Springer, LNAI.

——— (2009b): "Ultimate Cognition à la Gödel," *Cognitive Computation*, 1, 177–193.

Schwab, K. (2016): *The Fourth Industrial Revolution*. World Economic Forum.

Silver, D., T. Hubert, J. Schrittwieser, et al. (2018): "A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go Through Self-Play," *Science*, 362, 1140–1144.

Solow, R. M. (1956): "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 65–94.

——— (1957): "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39, 312–320.

Tegmark, M. (2017): *Life 3.0: Being Human in the Age of Artificial Intelligence.* New York: Knopf.

Temple, J. (2006): "Aggregate Production Functions and Growth Economics," *International Review of Applied Economics*, 20, 301–317.

Timmer, M. P., and B. van Ark (2005): "IT in the European Union: A Driver of Productivity Divergence?," *Oxford Economic Papers*, 57(4), 693–716.

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

Yudkowsky, E. (2013): "Intelligence Explosion Microeconomics," Technical report 2013-1, Machine Intelligence Research Institute.