NET WORK LISER

Discussion Paper Series

IZA DP No. 19110

November 2026

From Automation to Acceleration: Transitional Regimes in AI-Driven Growth

Wim Naudé
RWTH Aachen University,
University of Coimbra
and IZA@LISER

Tiago Neves Sequeira
University of Coimbra

Derick Almeida
University of Coimbra

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA@LISER Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.

ISSN: 2365-9793

IZA DP No. 19110 November 2026

From Automation to Acceleration: Transitional Regimes in AI-Driven Growth

Abstract

Can artificial intelligence (AI) turn automation into a sustained acceleration of economic growth? This paper argues that the answer depends on the transitional dynamics: when acceleration starts, what drives it, and when it ends. It develops a semi-endogenous growth model of the multi-stage transitional dynamics of AI-driven automation in scientific research. The novelty of the model is to separate what AI can do in principle (the technical automation frontier) from how much research output actually responds to AI (the effective elasticity), and both from whether growth actually accelerates due to transformative AI (TAI). It shows that the date at which AI starts to matter for research is the outcome of a race: growing knowledge widens the set of research tasks that AI can perform, which loosens the human bottleneck, while faster accumulation of AI capital adds AI input faster than tasks open up, which tightens it. Calibrated to historical U.S. task-automation shares and hardware energy-efficiency trends, four scenarios show that the same economy can experience no acceleration (declining growth despite rising automation), a modest transitory acceleration, or an explosive surge that is ultimately curbed by energy limits. Under strong task complementarity, early acceleration is driven by task reallocation rather than by AI-capital accumulation. Near-complete automation is neither necessary nor sufficient for permanent explosive growth; the outcome depends on the substitutability of AI and human researchers, and on how hard ideas are to find.

JEL classification

O31, O33, O40, O41

Keywords

artificial intelligence, automation, economic growth, ideas production function, transformative AI, transitional dynamics

Corresponding author

Wim Naudé

naude@time.rwth-aachen.de

1 Introduction

Progress in artificial intelligence (AI) has revived an old question in growth theory: can technological progress become sufficiently self-reinforcing to produce a sustained acceleration in economic growth?[1] AI systems can generate and evaluate hypotheses, design experiments and automate parts of research. If these capabilities raise the rate at which ideas are produced, AI may raise not only current output but also future production possibilities. The mechanism is a feedback loop: more knowledge raises output; higher output finances computing systems, models and robotic laboratories; and a larger stock of AI capital supports further research.

Whether this loop produces acceleration is not assured. Automated and human-performed research tasks may be complements, so that the remaining human tasks become bottlenecks (Aghion et al., 2019; Jones and Tonetti, 2026). Additional research effort may be duplicated, and ideas may become harder to find as knowledge accumulates (Bloom et al., 2020). Both forces can neutralize rapid growth in AI-supported research. The literature has therefore focused on thresholds: the degree of automation at which a balanced-growth path ceases to exist, or the elasticity of substitution above which growth explodes (Aghion et al., 2019; Jones, 2022b, 2025; Trammell and Korinek, 2020).

This paper's contribution is to shift the focus from terminal states to the transitional path. A threshold describes where an economy ends up if it stays on one growth regime forever – a terminal state; it does not describe the regimes through which a real economy passes, in what order, and how long each lasts. Some recent contributions do simulate multi-stage adjustment paths: Jones and Tonetti (2026) calibrate an endogenous task-automation frontier to U.S. data and show how weak links delay a growth explosion, Korinek and Suh (2024) and Almeida et al. (2024) analyze multi-phase trajectories and temporary overshoots, and Erdil et al. (2025) combine semi-endogenous R&D, compute-driven automation financed from output, and hardware and energy limits in an integrated assessment model.

What remains less developed is an account of how the effective contribution of AI to research evolves along the path: when acceleration starts, what drives it at each stage, when the human bottleneck is passed, and what ends the acceleration. Hence, this paper develops a semi-endogenous growth model to study the impact of AI on economic growth as a path. It

[1] Some of the discussion concerns *explosive* growth, defined as growth in gross world product of 30 per cent per year or more (Davidson, 2021; Erdil and Besiroglu, 2023; Trammell and Korinek, 2020). In this paper the term will refer either to this explicit numerical meaning, or to the mathematical meaning, which is a growth rate that rises without bound.

makes three contributions. First, it introduces an analytical separation between the technical automation frontier and the effective AI elasticity in research, and derives an exact law of motion for the latter wherein overcoming the human bottleneck is modelled as a dynamic race between endogenous, knowledge-driven frontier closure and the build-out of AI capital. Second, by decomposing the growth of research effort into AI accumulation, human research and task reallocation, it provides a two-stage mechanism: under gross complementarity, rapid initial AI investment temporarily depresses the effective AI elasticity, while early acceleration is driven by task reallocation rather than by AI-capital accumulation. Third, the quantitative calibration, to historical U.S. task-automation shares and hardware energy-efficiency trends, bridges the divide between macro-skeptics and techno-optimists: variations in fishing out, duplication and the substitution elasticity determine whether an economy experiences no acceleration, a modest transitional boom, or an explosive surge that is ultimately curbed by physical energy limits.

The paper is related to the literature that shows how task structure and bottlenecks shape the macroeconomic effects of AI (Aghion et al., 2019; Acemoglu, 2025; Jones, 2025; Jones and Tonetti, 2026), to the literature on feedback and explosive growth (Davidson et al., 2026; Erdil and Besiroglu, 2023; Trammell and Korinek, 2020; Nordhaus, 2021); and more generally to the literature on modelling AI in economics (Agrawal et al., 2019; Erdil et al., 2025; Naudé et al., 2024; Gries and Naudé, 2022).

The main findings of this paper are as follows. First, under strong task complementarity, what AI can do in principle is not the same as how much it actually helps research: as long as some research tasks still need humans, those tasks hold everything back.

Second, under the same condition, the effective AI elasticity, which measures how much research output actually responds to AI capital, does not rise gradually. It stays close to zero for decades and then rises sharply once the remaining human bottleneck is passed. The jump takes about ten years in the weak fishing-out scenarios; in the baseline the transition is more gradual, unfolding over roughly 25–30 years around 2086, and under strong duplication overcoming the human bottleneck is delayed until 2258.

Third, again under complementarity, the engine of growth changes over time. In the early stage, growth speeds up mainly because tasks are handed from humans to machines, while rapid AI investment temporarily depresses the effective AI elasticity. Only after the human bottleneck is overcome does building more AI capital become the main driver of faster growth in research effort, and through it of faster idea production and economic growth. Fourth, the outcome depends less on how much work is automated than on how easily AI

can replace human researchers, and how hard it is to find new ideas. Depending on these, the same economy can see no acceleration (declining growth despite rising automation), a modest transitory acceleration, or an explosive surge. In the fastest cases the energy ceiling is reached within a few decades; it terminates the acceleration and steps growth down to an efficiency-constrained rate (about 16 per cent per year in the calibration), rather than ending growth altogether, unless efficiency gains taper to zero.

Section 2 provides a brief overview of the related literature. Section 3 presents the model. Section 4 describes the growth regimes and the transitions between them. Section 5 calibrates the automation frontier and describes the various scenarios to be simulated. Section 6 reports the simulation results. Section 7 concludes. The proofs of all propositions are included in Appendix A.

2 Related Literature

The main strands of related literature deal with ideas and semi-endogenous growth, and with automation and AI in research. Here these are briefly reviewed, and the specific gap in this literature that is filled in this paper is identified and motivated.

The key insight on which much of the literature on technological innovation and growth is based, is that ideas are non-rival and are produced by researchers (Romer, 1990; Jones, 2019). In the semi-endogenous tradition the ideas production function is $\dot{A}_t = \theta S_t^{\lambda} A_t^{\varphi}$, where S_t is research effort, λ governs duplication ("stepping on toes") and $\varphi = 1 - \beta$ governs the intertemporal spillover (Jones, 1995, 2022b). If $\beta > 0$ ideas get harder to find, and long-run growth equals $\lambda n/\beta$ with n the growth of researchers. Bloom et al. (2020) estimate $\beta \approx 3.1$ for the U.S. aggregate economy with λ normalized to one, and $\beta \approx 0.2$ for Moore's Law. Klüppel and Knott (2023) however report firm-level evidence that points to $\varphi > 1$ and attribute falling aggregate research productivity to pathologies in the organization of research. These empirical findings partly underpin arguments that innovation is declining, and that with population growth rates slowing down (hence reducing the number of researchers in S_t), economic growth would ultimately peter out (Jones, 2022a).

On automation in research, the question is whether AI, as an innovation in the method of innovation (Agrawal et al., 2023), can help overcome these headwinds to innovation and growth. Aghion et al. (2019) model ideas production as a CES aggregate of tasks, some automated, following the task approach of Zeira (1998) and Acemoglu and Restrepo (2019).

With complementary tasks, hard-to-automate tasks bind and growth need not explode even under near-complete automation. Jones (2025) shows that the effect of AI on research depends jointly on task coverage, AI productivity and bottlenecks, and Jones and Tonetti (2026) estimate the share of private-business tasks assigned to capital in the United States and show how weak links delay a growth explosion; their endogenous task-automation frontier is calibrated to 1950–2023 U.S. data and simulated forward over several decades. Acemoglu (2025) emphasizes the limited gains from automating easy-to-learn tasks. Models of AI as an accumulable substitute for labor generate explosive growth in the absence of other constraints (Davidson, 2021; Erdil and Besiroglu, 2023; Trammell and Korinek, 2020; Korinek and Suh, 2024), and Nordhaus (2021) finds little evidence of an approaching singularity. Almeida et al. (2024) give AI knowledge its own law of motion and find that, for plausible parameters, explosive dynamics are at most a short-lived transitional overshoot, and Korinek and Suh (2024) analyze multi-phase trajectories. Erdil et al. (2025) build an integrated assessment model (GATE) that combines semi-endogenous R&D, task automation driven by compute accumulated from output, and the hardware-efficiency limits discussed by Ho et al. (2023) and Erdil and Besiroglu (2023). Evidence on the organization of research suggests AI may also create coordination and verification burdens that lower the effectiveness of added research effort (Ribeiro et al., 2023; Ball, 2023).

Several of the contributions discussed in the preceding paragraph therefore already trace multi-stage transitional trajectories in calibrated simulations (Jones and Tonetti, 2026; Korinek and Suh, 2024; Almeida et al., 2024; Erdil et al., 2025), and this paper does not claim that the transitional path has been ignored. Its contribution is more specific and analytical. First, it derives an exact differential law of motion for the effective AI elasticity $\varepsilon_{C,t}$ (Proposition 2), which characterizes a dynamic race between knowledge-driven frontier closure ($\chi A_t g_{A,t}$) and AI-capital build-out ($g_{C,t} - n$), the latter initially depressing $\varepsilon_{C,t}$ under gross complementarity. Second, it decomposes research-effort growth $g_{X,t}$ (Equation (8)) into AI accumulation, human research and task reallocation $\Gamma_t \dot{\alpha}_t$, to show why early transitional acceleration is driven by task reallocation rather than by AI-capital accumulation. The focus in the rest of the paper is therefore on the passage *between* growth regimes when the automation frontier is itself endogenous, AI capital is accumulated from output, and a physical energy ceiling bounds the long run.

3 The Model

This section sets out a growth model in which ideas drive long-run growth and are produced by research. It belongs to the family of semi-endogenous growth models (Jones, 1995, 2022b), in which ideas become harder to find as knowledge accumulates (Bloom et al., 2020). Following the task-based approach of Aghion et al. (2019), research combines AI capital and human researchers, and the two can be poor substitutes, so that tasks still done by humans may hold back the whole process.

The share of research tasks that AI can perform is not fixed: it expands as knowledge grows. AI capital is built up out of output, and a physical limit on energy use ends the process in the long run.

3.1 Production and research

Final output is produced with ideas, physical capital and labor:

$$Y_t = A_t K_t^{\kappa} L_{Y,t}^{1-\kappa}, \qquad 0 \leq \kappa < 1. \tag{1}$$

Population grows exogenously, $L_t = L_0 e^{nt}$. A constant share s_H of the population does research, $H_t = s_H L_t$, and $L_{Y,t} = (1 - s_H)L_t$. Physical capital accumulates according to $\dot{K}_t = s_K Y_t - \delta_K K_t$. AI capital enters the production of ideas but not final goods, which isolates the extent to which an accumulable research input can replace human researchers.

Aggregate research effort, X_t, is a CES aggregate of AI capital C_t (compute, models, autonomous agents, robotic laboratories and complements) and human research labor H_t:

$$X_t = \left[\alpha_t^{1-\rho} C_t^{\rho} + (1 - \alpha_t)^{1-\rho} H_t^{\rho}\right]^{1/\rho}, \qquad \rho < 1. \tag{2}$$

The elasticity of substitution is $\sigma = 1/(1 - \rho)$. Values $\rho < 0$ imply $\sigma < 1$ and gross complementarity, so that remaining human tasks are weak links; $0 < \rho < 1$ implies $\sigma > 1$. The weight $\alpha_t \in (0, 1)$ is the automation frontier. Ideas evolve according to

$$\dot{A}_t = \theta X_t^{\lambda} A_t^{1-\beta}, \qquad \theta, \lambda, \beta > 0, \tag{3}$$

so that the growth rate of knowledge, $g_{A,t} \equiv \dot{A}_t/A_t = \theta X_t^{\lambda} A_t^{-\beta}$, satisfies

$$\frac{\dot{g}_{A,t}}{g_{A,t}} = \lambda g_{X,t} - \beta g_{A,t} \equiv \mathcal{S}_t. \tag{4}$$

We call $\mathcal{S}_t$ the *knowledge-acceleration index*. Knowledge growth accelerates when $\mathcal{S}_t > 0$ and decelerates when $\mathcal{S}_t < 0$. The equality $\beta g_A = \lambda g_X$ characterizes constant growth only.

3.2 The endogenous automation frontier

With an initial knowledge stock A_0 and automation share $\alpha_0 \in [0, 1)$, the frontier evolves with knowledge:

$$\alpha_t = 1 - (1 - \alpha_0) \exp\left[-\chi(A_t - A_0)\right], \qquad \chi > 0. \tag{5}$$

Then $\alpha(A_0) = \alpha_0$, $\alpha_t \in [\alpha_0, 1)$ for $A_t \geq A_0$, $\alpha_t \to 1$ only as $A_t \to \infty$, and

$$\frac{\partial \alpha_t}{\partial A_t} = \chi(1 - \alpha_t) > 0, \qquad \frac{\partial^2 \alpha_t}{\partial A_t^2} = -\chi^2(1 - \alpha_t) < 0, \qquad \dot{\alpha}_t = \chi(1 - \alpha_t) A_t g_{A,t}. \tag{6}$$

Knowledge expands the frontier, but the same absolute increase in knowledge buys less automation as the frontier approaches one. Complete automation is approached asymptotically. The frontier (α_t) is a statement about what is technically possible. It need not coincide with how much research effort responds to AI.

3.3 Effective AI contribution and research-effort growth

Define the *effective AI elasticity* as the elasticity of research effort with respect to AI capital:

$$\varepsilon_{C,t} \equiv \frac{\partial \ln X_t}{\partial \ln C_t} = \left[1 + \left(\frac{1 - \alpha_t}{\alpha_t}\right)^{1-\rho} \left(\frac{H_t}{C_t}\right)^{\rho}\right]^{-1}, \qquad \varepsilon_{H,t} = 1 - \varepsilon_{C,t}. \tag{7}$$

It depends on the frontier, on the relative abundance of AI capital and human researchers, and on ρ. Under Cobb–Douglas aggregation ($\rho \to 0$) it equals α_t. Because α_t moves, log-differentiating (2) gives

$$g_{X,t} = \underbrace{\varepsilon_{C,t} g_{C,t}}_{\text{AI accumulation}} + \underbrace{(1 - \varepsilon_{C,t}) n}_{\text{human research}} + \underbrace{\Gamma_t \dot{\alpha}_t}_{\text{task reallocation}}, \qquad \Gamma_t \equiv \frac{\partial \ln X_t}{\partial \alpha_t} = \frac{1 - \rho}{\rho} \frac{\varepsilon_{C,t} - \alpha_t}{\alpha_t(1 - \alpha_t)}. \tag{8}$$

The sign of the reallocation coefficient Γ_t is the sign of $\rho(\varepsilon_{C,t} - \alpha_t)$. Under complementarity ($\rho < 0$) it is positive when the effective share is below the technical share, $\varepsilon_{C,t} < \alpha_t$, that is, when machines are abundant relative to the tasks they are allowed to perform and moving the marginal task to them relaxes the bottleneck. It is not simply that machines are abundant. It is that the effective AI elasticity is below the technical automation frontier, so that the expansion of the frontier increases the effective contribution of AI.

Under Cobb–Douglas aggregation $\Gamma_t \to \ln\{[C_t/\alpha_t]/[H_t/(1-\alpha_t)]\}$, the log ratio of AI to human input *per task*. Substituting (6) into (8) cancels the factor $(1-\alpha_t)$ in Γ_t, so that the reallocation term is $\Gamma_t \dot{\alpha}_t = \frac{1-\rho}{\rho} \frac{\varepsilon_{C,t} - \alpha_t}{\alpha_t} \chi A_t g_{A,t}$. Inserting the result into (4) gives the complete transition equation,

$$\mathcal{S}_t = \lambda \left[\varepsilon_{C,t} g_{C,t} + (1 - \varepsilon_{C,t}) n + \Gamma_t \chi (1 - \alpha_t) A_t g_{A,t} \right] - \beta g_{A,t}. \tag{9}$$

3.4 AI capital and the terminal constraint

AI capital is accumulated from output, $\dot{C}_t = s_C Y_t - \delta_C C_t$, so that $g_{C,t} = s_C Y_t / C_t - \delta_C$. The economy contains two feedbacks: $A_t \to \alpha_t \to X_t \to \dot{A}_t$ and $A_t \to Y_t \to C_t \to X_t \to \dot{A}_t$.

Operating AI capital requires energy, $E_t = e_t C_t$, where energy intensity falls at the rate $d_t \equiv -\dot{e}_t / e_t \geq 0$. Total energy use is bounded by a ceiling, $E_t \leq \bar{E}$. Define energy slack $\mathcal{E}_t \equiv \bar{E} - E_t$. While slack is positive, $g_{E,t} = g_{C,t} - d_t$. The ceiling first binds at

$$T_E \equiv \inf\{t \geq 0 : e_t C_t = \bar{E}\}, \qquad \text{which satisfies} \qquad \int_0^{T_E} (g_{C,s} - d_s)\, ds = \ln \frac{\bar{E}}{E_0}. \tag{10}$$

After T_E, $C_t = \bar{E}/e_t$ and $g_{C,t} = d_t$. Energy does not enter production directly. It limits the stock of AI capital that can be operated. In this paper the ceiling is a single parameter and a terminal regime.

Because s_C is a constant share of output, the investment stream $s_C Y_t$ continues after T_E although operable AI capital can only grow at the rate d_t. Maintaining the capped stock requires investment of $(\delta_C + d_t) C_t$; the surplus, $s_C Y_t - (\delta_C + d_t) C_t$, buys capacity that cannot be operated. We treat this surplus as stranded spending: it adds neither to operable AI capital C_t nor to physical capital K_t, which continues to accumulate from $s_K Y_t$. Allowing the surplus to be redirected, or allowing the ceiling itself to rise through price-driven extensions of electricity capacity, would soften the terminal regime.

4 Growth Regimes and Transitions

Because population and the reproducible stocks grow, the relevant steady states are steady states in growth rates. A *balanced-growth path* (BGP) is a path on which the principal aggregates grow at constant rates and $\varepsilon_{C,t}$ converges to a constant (see the remark after Proposition 1 on when this is attainable). *Transitional acceleration* means that $g_{A,t}$ rises along the adjustment path, whether or not the economy subsequently converges to a finite BGP. The growth regimes are defined through the two state variables $\mathcal{S}_t$ and $\mathcal{E}_t$: (i) *partial AI*: $\mathcal{S}_t \leq 0$ and $\mathcal{E}_t > 0$; (ii) *transformative AI* (TAI): $\mathcal{S}_t > 0$ and $\mathcal{E}_t > 0$; (iii) *energy-constrained AI*: $\mathcal{E}_t = 0$.

These definitions partition the state space without changing either technology. The first acceleration date is $T_A \equiv \inf\{t : \mathcal{S}_t > 0\}$, and the *bottleneck-release date* is

$$T_B \equiv \inf\{t \geq 0 : \varepsilon_{C,t} \geq \tfrac{1}{2}\}. \tag{11}$$

The pointwise regime $\mathcal{S}_t > 0$, entered at T_A, must be distinguished from *persistent* TAI, defined in Section 4.3, which requires acceleration to last until the ceiling binds. The date T_B is not part of the TAI definition. It identifies when additional AI capital, rather than additional human researchers, becomes the main way of expanding research effort.

4.1 Balanced growth under partial AI

Suppose the ceiling is slack, $\varepsilon_{C,t} \to \bar{\varepsilon}_C$, $\Gamma_t \dot{\alpha}_t \to 0$, and capital-output ratios converge. Along such a path $g_C = g_K = g_Y = n + g_A/(1-\kappa)$.

Proposition 1 (Partial-AI balanced growth) *A finite positive balanced-growth rate exists if and only if*

$$\bar{\varepsilon}_C < \varepsilon_C^* \equiv \frac{\beta(1-\kappa)}{\lambda}. \tag{12}$$

Then $g_A^P = \lambda n/[\beta - \lambda\bar{\varepsilon}_C/(1-\kappa)]$ *and* $g_Y^P = n + g_A^P/(1-\kappa)$*, and* g_A^P *is locally asymptotically stable conditional on* $\bar{\varepsilon}_C$ *and on converged capital–output ratios. As* $\bar{\varepsilon}_C \uparrow \varepsilon_C^*$ *the balanced-growth rate becomes arbitrarily large.*

The proposition generalizes the semi-endogenous result $g_A = \lambda n/\beta$. Population growth remains the ultimate source of research-input growth, but AI capital amplifies it through $\bar{\varepsilon}_C$.

Two qualifications need to be raised at this point. First, the threshold concerns the *effective* elasticity: in the Cobb–Douglas case it becomes $\alpha^* = \beta(1-\kappa)/\lambda$, which exceeds one under the aggregate calibration $\beta = 3.1$, $\lambda = 1$. Second, the non-existence of a BGP does not by itself imply acceleration or finite-time explosion. Acceleration depends on the sign of $\mathcal{S}_t$ along the transition.

Because the frontier (5) depends on the absolute level of knowledge A_t rather than on proportional changes $\ln(A_t/A_0)$, any path with strictly positive constant knowledge growth makes A_t grow exponentially, so that $\chi A_t g_{A,t}$ diverges and $\alpha_t \to 1$. If $g_{C,t}$ converges to a finite constant, $\ln(C_t/H_t)$ grows only linearly while $(1-\rho)\chi A_t$ grows exponentially, so that $\Xi_t \to -\infty$ in Proposition 2 and $\varepsilon_{C,t} \to 1$. A hypothesised interior limit $\bar{\varepsilon}_C < 1$ with $\alpha_t \to 1$ would also make the reallocation term $\Gamma_t \dot{\alpha}_t$ diverge rather than vanish. An interior $\bar{\varepsilon}_C \in (0,1)$ in Propositions 1 and 5 should therefore be read as (i) the fixed-frontier benchmark ($\chi = 0$), (ii) a quasi-steady-state approximation, valid while α_t and $\varepsilon_{C,t}$ move slowly, or (iii) the outcome under a log-knowledge frontier, $\alpha_t = 1 - (1-\alpha_0)(A_t/A_0)^{-\chi}$, in which only proportional changes in knowledge matter. In the true asymptotic limit with $\chi > 0$, any balanced-growth path with positive knowledge growth has $\bar{\varepsilon}_C = 1$. It then exists if and only if $1 < \varepsilon_C^*$, that is $\beta(1-\kappa) > \lambda$, with $g_A^P = \lambda n/[\beta - \lambda/(1-\kappa)]$, which equals 0.64 per cent per year in the Bloom et al. calibration. After the ceiling binds, the same logic gives $\bar{\varepsilon}_C = 1$ and $g_A^E = \lambda d/\beta$ (Proposition 5).

4.2 Weak links and the dynamics of the effective AI elasticity

At a fixed frontier, the limit of $\varepsilon_{C,t}$ as AI capital becomes abundant ($C_t/H_t \to \infty$) is 0 if $\rho < 0$, α_t if $\rho \to 0$ (the fixed frontier level), and 1 if $0 < \rho < 1$. This is the weak-link result: under complementarity, abundant AI capital cannot remove the human bottleneck while the frontier is incomplete. But the frontier is not fixed, and the object of interest is the *path* of $\varepsilon_{C,t}$.

Proposition 2 (Dynamics of the effective AI elasticity) *The effective AI elasticity satisfies the exact law of motion*

$$\dot{\varepsilon}_{C,t} = \varepsilon_{C,t}(1-\varepsilon_{C,t}) \left[\frac{(1-\rho)}{\alpha_t} \chi A_t g_{A,t} - (-\rho)\,(g_{C,t} - n) \right]. \tag{13}$$

Let $\rho < 0$ and $A_t \to \infty$. Define $\Xi_t \equiv (-\rho)\ln(C_t/H_t) - (1-\rho)\chi A_t$. (i) If $\Xi_t \to -\infty$ then $\varepsilon_{C,t} \to 1$; (ii) if $\Xi_t \to +\infty$ then $\varepsilon_{C,t} \to 0$; (iii) if $\Xi_t \to \bar{\Xi}$ then $\varepsilon_{C,t} \to [1 +$

$e^{\bar{\Xi}+(1-\rho)(\ln(1-\alpha_0)+\chi A_0)}\big]^{-1}$.

Here, it is assumed that $0 < \alpha_t < 1$, $C_t > 0$, and $H_t > 0$.

Equation (13) shows a *race*. The first term is the closing of the frontier, which raises the effective elasticity. The second is the accumulation of AI capital faster than researchers, which lowers it under complementarity. The elasticity rises if and only if

$$(1-\rho)\,\chi A_t g_{A,t} \;>\; \alpha_t\,(-\rho)\,(g_{C,t}-n). \tag{14}$$

The condition (14) moves the analysis beyond a threshold analysis. It involves the speed of knowledge growth, which is endogenous, and the speed of AI-capital accumulation, which depends on investment. A fast AI build-out (g_C high) *delays* bottleneck release under complementarity, whereas faster knowledge growth *advances* it. The transition from a near-zero to a near-one effective elasticity is also sharp, because $\varepsilon_{C,t}(1-\varepsilon_{C,t})$ in (13) is largest in the middle: once the race tips, the elasticity moves quickly.

4.3 Transitional acceleration and persistent transformation

Proposition 3 (Transitional acceleration under partial AI) *While the ceiling is slack, knowledge growth accelerates if and only if*

$$\lambda\left[\varepsilon_{C,t} g_{C,t} + (1-\varepsilon_{C,t})n + \Gamma_t \dot{\alpha}_t\right] > \beta g_{A,t}. \tag{15}$$

Suppose the BGP of Proposition 1 exists. Acceleration can still occur when AI capital initially grows rapidly (it grows fast in percentage terms from a small base) or when the reallocation term is sufficiently positive. It is temporary if the left-hand side of (15) converges to βg_A^P from above.

A transitional boom therefore must be distinguished from transformative AI. We say that AI is persistently transformative from date t if $\mathcal{S}_s > 0$ for all $s \in (t, T_E)$, and define

$$T_{\text{TAI}} \equiv \inf\{t \in [0, T_E) : \mathcal{S}_s > 0 \text{ and } \mathcal{E}_s > 0 \text{ for all } s \in (t, T_E)\}, \qquad \inf \emptyset = +\infty. \tag{16}$$

Restricting candidate dates to $[0, T_E)$ avoids the vacuous truth of the condition for all $t \geq T_E$, and the open interval allows $\mathcal{S}_{t_0} = 0$ at the crossing date itself. A path with $\mathcal{S}_t > 0$ on a

bounded interval that returns to $\mathcal{S}_t \leq 0$ before T_E (or with $T_E = \infty$) has $T_{\text{TAI}} = \infty$: it is a transitional acceleration, not persistent TAI.

Proposition 4 (Persistent transition to transformative AI) *Suppose the ceiling is slack on an interval beginning at t_0, $\mathcal{S}_{t_0} = 0$, and the crossing is transversal, $\dot{\mathcal{S}}_{t_0} > 0$. Then knowledge growth accelerates immediately after t_0. If $\mathcal{S}_t$ does not return to zero before T_E, then $t_0 = T_{\text{TAI}}$.*

The proposition states a persistence condition rather than assuming that the first crossing is irreversible. The economy can cross back into partial AI if AI-capital growth slows, the reallocation effect fades or the human bottleneck tightens.

4.4 The terminal regime and the ordering of regimes

Proposition 5 (Terminal energy regime) *Suppose the ceiling is initially slack, and let $g^u_{C,t} \equiv s_C Y_t/C_t - \delta_C$ be the growth rate of AI capital that the investment rule would deliver if the ceiling were ignored (the unconstrained path). The ceiling binds at a finite date if $\int_0^\infty (g^u_{C,s} - d_s)ds > \ln(\bar{E}/E_0)$, in particular if $g^u_{C,t} - d_t \geq \eta > 0$ eventually. Once reached, it remains binding for all $t \geq T_E$ if and only if $s_C Y_t/(\bar{E}/e_t) - \delta_C \geq d_t$ for all $t \geq T_E$. If, after T_E, $d_t \to d \geq 0$, $\varepsilon_{C,t} \to \bar{\varepsilon}_C$ and $\Gamma_t \dot{\alpha}_t \to 0$, then*

$$g^E_A = \frac{\lambda}{\beta}\left[\bar{\varepsilon}_C d + (1 - \bar{\varepsilon}_C)n\right], \qquad g^E_Y = n + \frac{g^E_A}{1-\kappa}, \tag{17}$$

and g_A converges globally on $g_A > 0$ to g^E_A whenever $g^E_A > 0$. With $\chi > 0$, $\bar{\varepsilon}_C = 1$ (see the remark in Section 4), so that $g^E_A = \lambda d/\beta$. If efficiency gains taper to zero ($d = 0$) and $\bar{\varepsilon}_C < 1$, $g^E_A = \lambda(1 - \bar{\varepsilon}_C)n/\beta$; under full research automation ($\bar{\varepsilon}_C = 1$) the right-hand side of (17) is zero, g_A converges to the boundary point 0 at rate $g_{A,t} \sim (\beta t)^{-1}$ (so that $A_t \sim \beta^{-1} \ln t \to \infty$ and $\bar{\varepsilon}_C = 1$ remains self-consistent), and $g^E_Y = n$.

The ceiling converts the accumulable research input into one whose growth is capped by efficiency improvement. Faster knowledge growth raises output and AI investment and so brings T_E forward, while efficiency improvement delays it. The levels of A, K, C and Y are continuous at T_E, with a kink in C_t. The growth rates $g_{C,t}$ and $g_{X,t}$ jump discontinuously at T_E. By contrast, $g_{A,t} = \theta X_t^\lambda A_t^{-\beta}$ and $g_{K,t} = s_K Y_t/K_t - \delta_K$ are continuous functions of the levels, so $g_{A,t}$, $g_{K,t}$ and $g_{Y,t}$ are continuous at T_E and only their time derivatives have a kink; this is why output growth can peak slightly after T_E.

Proposition 6 (Regime ordering) *Let $\mathcal{S}_0 \leq 0$ and $\mathcal{E}_0 > 0$. (a) If $T_{\text{TAI}} = \infty = T_E$, the economy may pass through episodes with $\mathcal{S}_t > 0$ (transitional acceleration, entered at T_A) but never becomes persistently transformative; it ends in partial AI and converges to the BGP of Proposition 1 (read as in the remark following that proposition). (b) If $T_{\text{TAI}} < T_E < \infty$ the economy passes from partial AI to transformative AI to energy-constrained AI and, under Proposition 5, converges to g_A^E. (c) If $T_E < \infty$ and $T_{\text{TAI}} = \infty$, that is, $\mathcal{S}_t$ does not stay positive up to T_E, it moves from partial AI (possibly after transitory episodes of acceleration) directly to energy-constrained partial AI. (d) If $T_{\text{TAI}} < \infty = T_E$, it enters TAI without reaching the ceiling.*

When $\mathcal{S}_0 > 0$, as in the scenarios below, the economy starts in the transformative region ($T_A = 0$). Whether this is persistent TAI depends on whether $\mathcal{S}_t$ stays positive until T_E: in the Moore's Law and AI Boom scenarios it does, and (b) applies from the initial date; in the Bloom et al. baseline $\mathcal{S}_t$ turns negative after about 66 years while the ceiling never binds, so $T_{\text{TAI}} = \infty$ and the episode is a transitory acceleration covered by (a). T_B is also used to date the internal stages of the acceleration region.

5 Calibration and Scenarios

The theoretical model developed in Sections 3 and 4 describes which growth regimes an AI-driven economy can pass through, but not how long each lasts, how large the changes are, or which assumptions matter most. This section answers those questions with simulations. The model starts in 2023 and is run for 300 years under four scenarios. These differ in how hard ideas are to find, how much new research effort adds duplicated work, and how fast AI capital is built. This section describes the calibration choices and the scenarios.

5.1 The automation frontier

The parameter χ is calibrated using the estimates of Jones and Tonetti (2026) of the share of U.S. private-business tasks assigned to capital, $\{\alpha_{1950}, \alpha_{1975}, \alpha_{2000}, \alpha_{2023}\} = \{0.28, 0.45, 0.63, 0.75\}$, and a TFP-based knowledge stock. Following Bloom et al. (2020), aggregate idea output is represented by TFP, which grew at approximately 1.5 per cent per year; we normalize $A_{1950} = 1$ and set $A_t = \exp[0.015(t - 1950)]$, so that $A = 1.000, 1.455, 2.117, 2.989$. Because the level of a TFP index has an arbitrary origin,

(5) is anchored at $\alpha_0 = 0.28$ and $A_0 = 1$.

Minimizing squared deviations between observed and model-implied shares,

$$\widehat{\chi} = \arg\min_{\chi>0} \sum_t \left\{\alpha_t - \left[1 - (1-\alpha_0)e^{-\chi(A_t - A_0)}\right]\right\}^2 = 0.565, \tag{18}$$

which gives $\alpha_t = 1 - 0.72\exp[-0.565(A_t - 1)]$. The fitted shares are 0.280, 0.443, 0.617 and 0.766, with a root mean squared error of 1.1 percentage points (See Figure 1a).

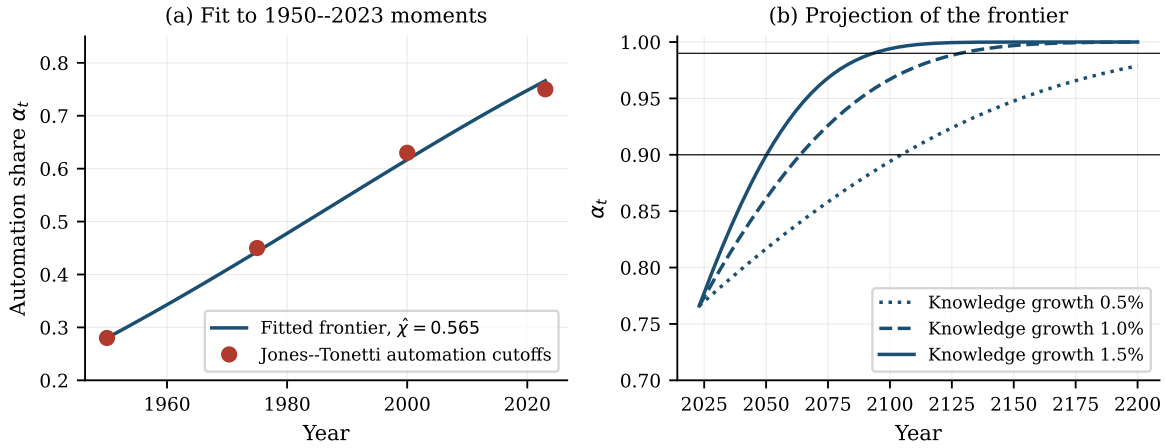


Figure 1: The calibrated automation frontier

Notes: Panel (a): fit to the four automation moments. Panel (b): projection of the frontier from 2023 for alternative constant growth rates of knowledge. Horizontal lines mark 0.90 and 0.99. Authors' calculations.

The scale of χ depends on the normalization of A: rescaling A by c rescales χ by $1/c$. Table 1 reports sensitivity to the assumed growth of knowledge. Faster measured knowledge growth requires a smaller χ to account for the same historical automation. The timing of the frontier is sensitive to this assumption: if knowledge continued to grow at 1.5 per cent per year the frontier would reach 0.90 around 2050, but at 0.5 per cent per year it would take until about 2105. In every case complete automation is approached only asymptotically.

The calibration is subject to two caveats. One is that the Jones and Tonetti (2026) cutoffs apply to private business as a whole, whereas (2) uses the automation share inside research. The baseline implicitly assumes that the historical relation between knowledge and task automation in private business is informative about research, and we therefore examine alternative values of χ below. Second, TFP is an imperfect measure of aggregate idea output.

Table 1: Sensitivity of the calibrated $\widehat{\chi}$ to the growth rate of knowledge

Annual growth of A_t	1.0%	1.2%	1.5%	2.0%
Calibrated $\widehat{\chi}$	0.995	0.778	0.565	0.357
RMSE	0.005	0.006	0.011	0.020
Year $\alpha_t = 0.90$ (post-2023 growth equal to calibration)	2059	2054	2050	2044

Notes: Each column normalizes $A_{1950} = 1$ and estimates χ by nonlinear least squares on the four moments. Projected years assume knowledge keeps growing at the calibration rate, starting from the fitted 2023 share of 0.766 (the simulations in Section 6 instead start from the observed share of 0.75).

5.2 Scenarios

The simulations start in 2023 with annual data and a horizon of 300 years. Table 2 lists the parameters. Macroeconomic parameters follow Jones and Tonetti (2026): $n = 1$ per cent, $\kappa = 0.35$, $s_K = 0.20$, $\delta_K = 0.05$ and $K_0/Y_0 = 2.5$, which implies initial physical-capital growth of $s_K/(K_0/Y_0) - \delta_K = 3.0$ per cent and initial output growth of 2.7 per cent, slightly above the balanced rate $n + g_{A,0}/(1-\kappa) \approx 2.5$ per cent. The knowledge stock retains its 1950 normalization so that $A_0 = 2.989$ in 2023, preserving the scale in which χ was estimated. The intercept of (5) is, however, re-anchored at the observed 2023 share, $\alpha_0 = 0.75$, rather than at the fitted value of 0.766 used for the projections in Figure 1b and Table 1, while $\widehat{\chi} = 0.565$ is retained. This raises the unautomated share $1 - \alpha_t$ by about 7 per cent at every knowledge level, so the simulated frontier starts about 1.6 percentage points below the fitted one. The simulated paths also differ from the projections in Figure 1b because knowledge growth is endogenous in the simulations. Anchoring at the observed share is the more conservative choice for the timing results. All simulations use fourth-order Runge–Kutta integration with a step of 0.02 years, including those behind Figure 4 and Table 4. All scenarios target initial knowledge growth of one per cent, with θ recalibrated in each.

The substitution elasticity is $\sigma = 0.2$ ($\rho = -4$), the weak-link calibration of Jones and Tonetti (2026). The initial effective AI elasticity is targeted at $\varepsilon_{C,0} = 0.10$ and H_0 is chosen residually, so that the economically relevant objects are the ratio of inputs and the implied elasticity. Initial AI capital is 1 per cent of output, an R&D-specific target informed by the infrastructure evidence in Wachter and Wachter (2026), Carpinelli et al. (2026) and Korinek and McKelvey (2026). AI capital depreciates at 20 per cent per year to reflect hardware obsolescence. The initial build-out rate $g_{C,0}$ is 50 per cent in the moderated case or 140 per cent in the AI-boom case, which implies investment shares $s_C = 0.007$ and 0.016 of output.

Initial energy use is one quarter of the ceiling, and energy-efficiency improvement falls from

15 to 2 per cent per year, $d_t = 0.02 + 0.13e^{-0.04t}$, reflecting rapid recent hardware gains tempered by system-level overheads (Rahman, 2024; International Energy Agency, 2025; IEEE, 2024; Ho et al., 2023).

Table 2: Calibration values and scenarios

Object	Interpretation	Value	Source or target
n, κ, s_K, δ_K	Population growth; capital elasticity; investment rate; depreciation	0.01; 0.35; 0.20; 0.05	Jones and Tonetti (2026)
A_0, $g_{A,0}$	Knowledge in 2023; initial knowledge growth	2.989; 0.01	TFP index; scenario target
α_0, χ	Initial automation; frontier response	0.75; 0.565	Jones and Tonetti (2026); fit to moments
σ (ρ)	Substitution elasticity	0.2 (-4)	Jones and Tonetti (2026)
$\varepsilon_{C,0}$, C_0/Y_0	Initial effective AI share; AI capital ratio	0.10; 0.01	Scenario targets
δ_C, $g_{C,0}$	AI depreciation; initial AI-capital growth	0.20; 0.50 or 1.40	Scenario; Korinek and McKelvey (2026)
$E_0/\bar{E}$	Initial use of energy ceiling	0.25	Scenario
d_0, d_∞, μ_d	Efficiency improvement: initial, long-run, speed	0.15; 0.02; 0.04	Rahman (2024); International Energy Agency (2025)

Scenarios: $(\lambda, \beta, g_{C,0})$ = Bloom et al. (1.0, 3.1, 0.5); Duplication (0.2, 3.1, 0.5); Moore's Law (1.0, 0.2, 0.5); AI Boom (1.0, 0.2, 1.4).

The first scenario uses the aggregate idea-production estimate of Bloom et al. (2020), $\lambda = 1$ and $\beta = 3.1$. The second isolates stronger duplication by lowering λ to 0.2 and leaving β unchanged. The third and fourth use the technology-specific estimate $\beta = 0.2$ for Moore's Law, with the moderated build-out and the AI-boom build-out respectively. The first two scenarios therefore differ in duplication, the first and third in fishing out, and the third and fourth in initial investment.

6 Quantitative Results

This section presents the simulation results for the four calibrated scenarios. The results can be summarised at the outset.

First, the outcome depends strongly on how hard ideas are to find. Under the aggregate estimate, growth speeds up only modestly, peaking at about 3.6 per cent a year in 2090, and when extra research effort is largely duplicated it does not speed up at all. Under the technology-specific estimate, growth reaches several tens of per cent a year (all growth rates

are instantaneous, continuous-time rates), with output growth peaking at about 96 per cent in the Moore's Law scenario and 83 per cent in the AI Boom, in the 2060s.

Second, under strong complementarity the road to the fast growth outcomes runs in stages. AI helps research very little for decades, then its contribution jumps within about ten years once the human bottleneck is passed (more gradually, over 25–30 years, in the baseline with $\beta = 3.1$), and the energy limit is reached soon after, which ends the acceleration and steps growth down to an efficiency-constrained rate.

Third, none of the four scenarios produces permanent explosive growth. Even in the fastest cases, growth is still about 25 per cent in 2123 and falls back to about 20 per cent a year a century after the energy limit binds, converging to about 16 per cent because efficiency improvement continues.

6.1 Four transitions

Figure 2 and Table 3 report the transitions.

Table 3: Summary of the simulated transitions

Scenario	Peak g_A (year)	Peak g_Y (year)	Bottleneck release T_B	Ceiling binds T_E	g_Y in 2123
Bloom et al.	1.8% (2089)	3.6% (2090)	2086	Does not bind	3.0%
Duplication	1.0% (2023)	2.7% (2023)	2258	Does not bind	1.5%
Moore's Law	76.7% (2065)	96.3% (2066)	2052	2065	25.8%
AI Boom	66.9% (2058)	83.4% (2060)	2053	2058	25.1%

Notes: T_B is the first year with $\varepsilon_{C,t} \geq 1/2$. The ceiling binds 41.9 years after 2023 in the Moore's Law scenario and 35.4 years after 2023 in the AI Boom. It does not bind within 300 years in the first two scenarios. Peak years are rounded to the nearest calendar year. Output growth is computed from the exact decomposition $g_Y = g_A + \kappa g_K + (1 - \kappa)n$. All growth rates are instantaneous (continuous-time) log-derivatives, not compounded annual percentage changes, which matters at high growth rates.

Automation without acceleration. In the Duplication scenario, knowledge growth declines from its initial value of one per cent; it is 0.32 per cent after 100 years and 0.21 per cent after 200 years, with output growth of 1.53 and 1.34 per cent. The automation frontier still approaches 0.92 by 2123. Reducing λ from one to 0.2 removes the transitional acceleration even though fishing out and AI-capital build-out are unchanged. Automation expands, but too little additional idea output results.

Transitory acceleration. In the Bloom et al. baseline, $\mathcal{S}_t > 0$ for about the first 66 years, so that knowledge growth rises from one per cent to a peak of about 1.8 per cent around

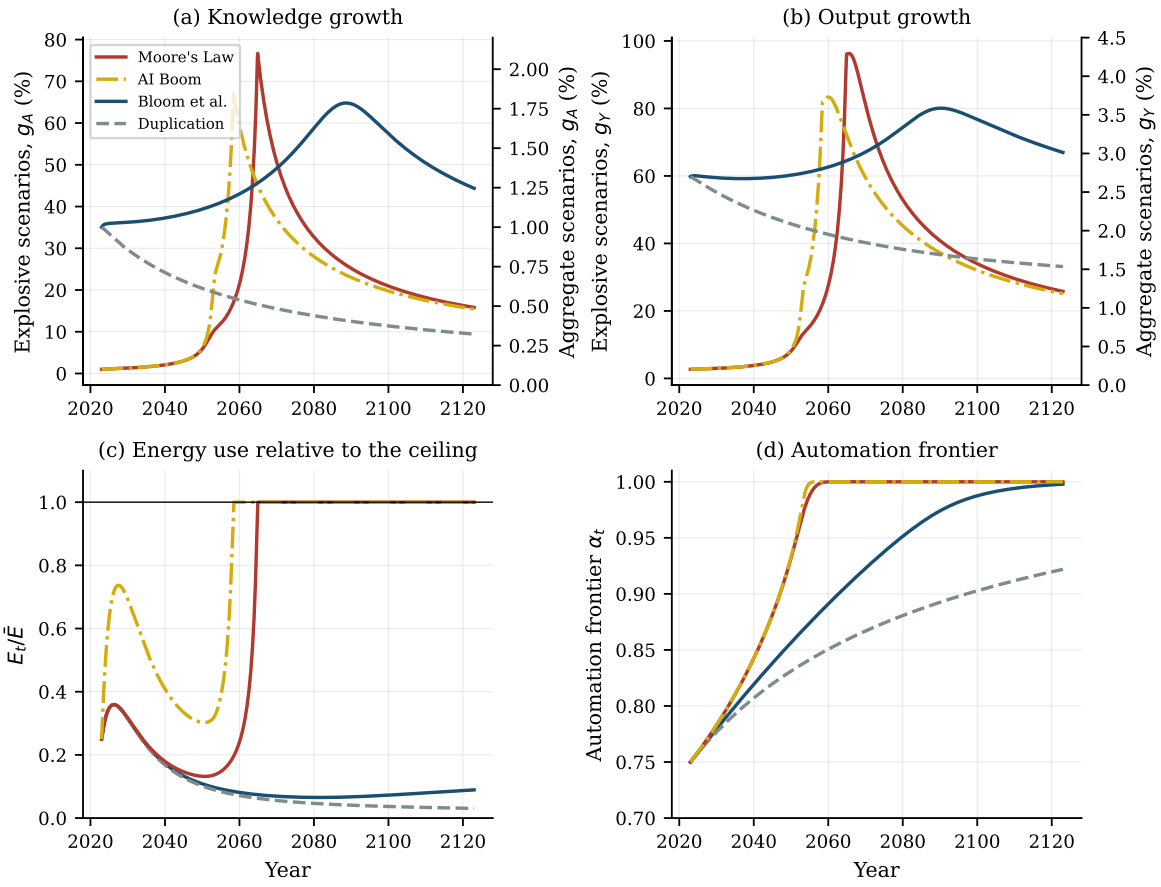


Figure 2: Growth, energy use and the automation frontier across scenarios

Notes: The upper panels use two vertical scales: left axes for the Moore's Law and AI Boom scenarios, right axes for the Bloom et al. and Duplication scenarios. The energy ceiling binds when $E_t/\bar{E} = 1$. Calendar time starts in 2023. Authors' simulations.

2089 and output growth peaks at 3.6 per cent in 2090. Both then decline: knowledge growth is 1.25 per cent after 100 years and 0.81 per cent after 200 years, and output growth 3.0 and 2.3 per cent. The frontier passes 0.95 in 2080 and approaches one, yet the economy returns toward a convergent path. Near-complete task automation is therefore not sufficient evidence of a permanent growth acceleration. The ceiling never binds, so $T_{\text{TAI}} = \infty$ in the sense of (16): the episode of acceleration is transitory and the economy returns to partial AI.

Explosive acceleration, then the ceiling. With weak fishing out ($\beta = 0.2$), knowledge growth reaches 76.7 per cent around 2065 in the Moore's Law scenario, with output growth peaking at 96.3 per cent. In the AI Boom the corresponding peaks are 66.9 and 83.4 per cent, but reached earlier, around 2058–2060. In both, knowledge growth peaks when the ceiling binds; output growth peaks at or shortly after that date (about two years later in the AI Boom, because physical-capital growth keeps rising briefly after T_E). *Faster AI investment advances the transition but lowers the peak*, because the ceiling intervenes while the explosive path is still lower. Growth then declines toward the energy-constrained rate: output growth is still about 25 per cent in 2123 (100 years after the 2023 start) and about 20 per cent 100 years after the ceiling binds. It converges to $g_Y^E = n + g_A^E/(1-\kappa)$ with $g_A^E = (\lambda/\beta)d_\infty = (1.0/0.2)(0.02) = 0.10$, that is $g_Y^E = 0.01 + 0.10/0.65 \approx 16.4$ per cent, because efficiency improvement remains at 2 per cent per year in this calibration (Proposition 5). With efficiency gains that taper to zero, the limit would fall toward n.

None of the four scenarios produces permanent explosive growth.

6.2 Analyzing a transition

Figure 3 depicts the transition. The top panels show that the technical frontier and the effective AI elasticity are very different objects. In the Moore's Law scenario, α_t rises smoothly from 0.75 to 0.95 between 2023 and 2051, while $\varepsilon_{C,t}$ falls from 0.10 to about 0.003, because AI capital grows faster than the frontier closes and (14) fails. Then, within about a decade, the elasticity rises from near zero to one. The bottleneck is passed in 2052 (T_B), 13 years before the energy ceiling binds. In the Bloom et al. baseline the sequence is the same but much slower, with T_B in 2086 and a transition that unfolds over roughly 25–30 years, and in the Duplication scenario T_B is delayed until 2258. These patterns hold under strong complementarity ($\sigma = 0.2$).

The bottom panels decompose the growth of effective research input. During the weak-link

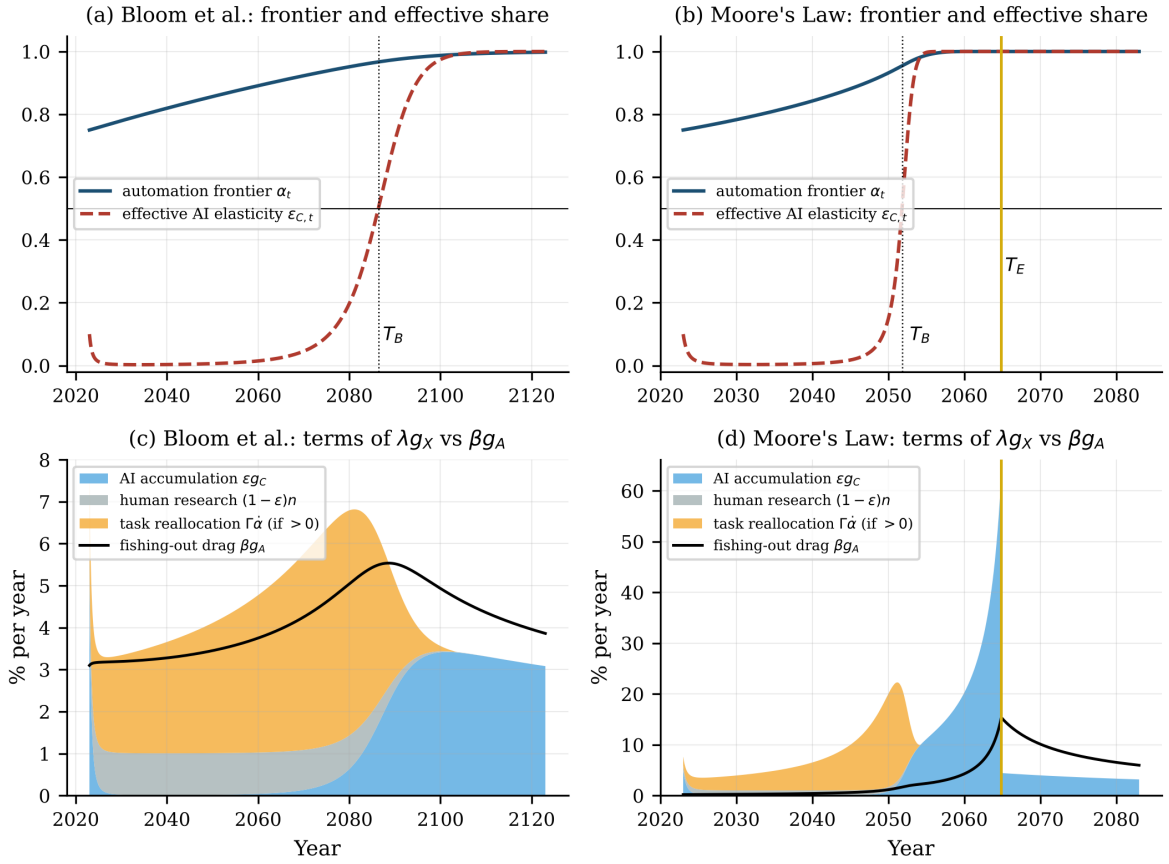


Figure 3: Analyzing a Transition

Notes: Top panels: automation frontier α_t and effective AI elasticity $\varepsilon_{C,t}$; dotted line: bottleneck release T_B; solid vertical line: ceiling binds T_E. Bottom panels: stacked terms of λg_X from equation (8) (the reallocation term is shown when positive) against the fishing-out drag βg_A; acceleration occurs when the stack exceeds the line. The reallocation term is computed from its closed form, $\Gamma_t \dot{\alpha}_t = \frac{1-\rho}{\rho} \frac{\varepsilon_{C,t}-\alpha_t}{\alpha_t} \chi A_t g_{A,t}$, not as a finite-difference residual, which would create a spurious spike at T_E. Authors' simulations.

phase, AI-capital accumulation contributes almost nothing, because $\varepsilon_{C,t} \approx 0$, and human research contributes λn. What lifts λg_X above the fishing-out drag is the reallocation term $\Gamma_t \dot{\alpha}_t$: in the Moore's Law scenario it rises from 1.7 per cent in 2023 to 7.2 per cent in 2043, and it is several times larger than the accumulation term, which starts at 5 per cent in 2023 but falls below 0.1 per cent within five years, while human research contributes $\lambda n \approx 1$ per cent; accumulation becomes important only shortly before the bottleneck is released. In the Bloom et al. baseline the reallocation term rises from about 1.7 to 5.0 per cent by 2073 and accounts for the acceleration there. Accumulation takes over only after T_B.

Under gross complementarity the transition has thus two stages with different causes: first the frontier-expansion channel, which is hump-shaped in knowledge and wanes as $\alpha_t \to 1$, and then AI-capital accumulation. A model with a fixed frontier, or one that attributes acceleration to the accumulation of AI capital, misses the first stage and mis-dates the second.

6.3 Critical parameters

Figure 4 maps the peak output growth over the substitution elasticity σ and the fishing-out parameter β, holding the Moore's Law build-out fixed. It shows that fishing out matters: for $\sigma \leq 0.5$, a value of β of 3.1 yields peaks below 7 per cent, and $\beta = 0.2$ yields peaks above 80 per cent.

It also indicates that the substitution elasticity causes a regime switch. At the aggregate value $\beta = 3.1$ the economy shows a modest transitional acceleration for all $\sigma \leq 0.5$, but peak growth rises to 10 per cent at $\sigma = 0.6$, 26 per cent at $\sigma = 0.7$ and above 280 per cent at $\sigma = 0.8$, with the energy ceiling binding within 19 years. Complementarity of the strength estimated by Jones and Tonetti (2026) ($\sigma = 0.2$) is therefore significant. An implication is that finding empirical evidence on σ in R&D is an avenue for further research.

Table 4 varies the frontier parameter χ and σ. A faster frontier (higher χ) advances the bottleneck release and the arrival of the ceiling and lowers the peak, because the ceiling intervenes earlier. Higher substitutability advances every date and raises the peak. The qualitative classification of the two scenarios, explosive for $\beta = 0.2$ and modest for $\beta = 3.1$, does not change across the grid, but the timing of release varies by a factor of three. The timing results are therefore the least robust output of the exercise and should be interpreted as illustrations only.

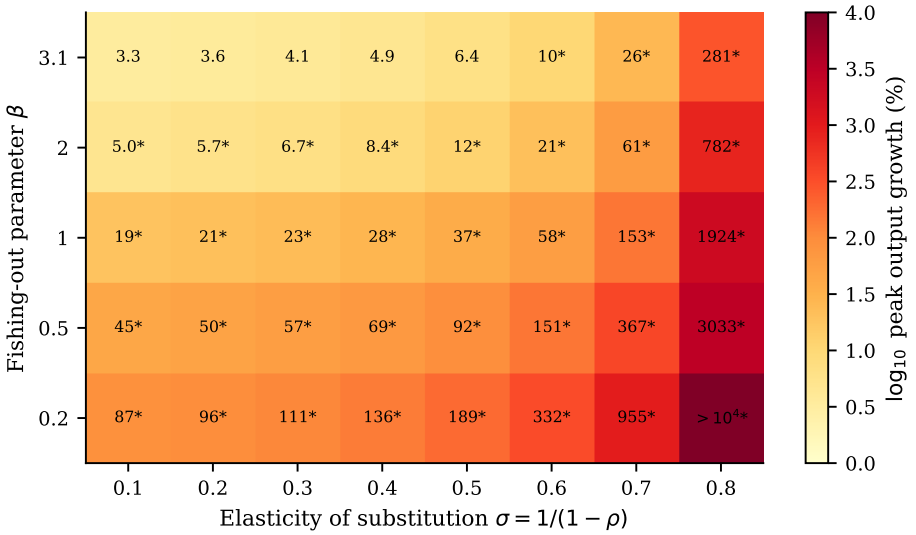


Figure 4: Growth regime map

Notes: Cell entries are peak output growth in per cent per year over a 300-year horizon, for $\lambda = 1$ and $g_{C,0} = 0.5$. An asterisk indicates that the energy ceiling binds. Colour shows $\log_{10}$ of peak growth. Integration step 0.02 years, as in Table 4. All entries represent instantaneous log growth rates expressed as percentages, not discrete annual percentage changes. Authors' simulations.

Table 4: Robustness of the transition dates to the frontier parameter and the substitution elasticity

		Moore's Law			Bloom et al.	
χ	σ	T_B (years)	T_E (years)	Peak g_Y	T_B (years)	Peak g_Y
0.357	0.1	43.2	53.7	108%	133.9	3.1%
0.357	0.2	40.2	49.4	122%	119.5	3.3%
0.357	0.5	30.3	34.4	262%	75.9	5.6%
0.565	0.1	30.8	46.2	87%	70.0	3.3%
0.565	0.2	28.8	41.9	96%	63.4	3.6%
0.565	0.5	21.8	27.1	190%	42.1	6.4%
0.995	0.1	19.0	40.4	73%	30.7	3.7%
0.995	0.2	18.0	35.7	78%	28.7	4.1%
0.995	0.5	13.9	20.5	134%	20.8	7.9%

Notes: Years are measured from 2023. $\chi = 0.565$ and $\sigma = 0.2$ is the baseline. In the Bloom et al. scenario the ceiling does not bind for any cell. χ is varied ceteris paribus, holding $A_0 = 2.989$ and $\alpha_0 = 0.75$ fixed, so the product χA_0 differs across rows. Integration step 0.02 years. Authors' simulations.

A further caveat concerns the initial automation share. Setting $\alpha_0 = 0.75$ in the research aggregator (2) literally implies that three quarters of research tasks were already automatable in 2023, leaving only a quarter of the frontier to be closed before the human bottleneck is released. Yet 0.75 is the private-business share of Jones and Tonetti (2026), not a research-specific estimate. A lower research-specific α_0 means a larger initial unautomated share $1 - \alpha_0$, so that frontier closure, which is what raises $\varepsilon_{C,t}$ in (13), takes longer. Holding other parameters fixed, one may therefore expect a later bottleneck release T_B and a later arrival of the ceiling.

7 Concluding Remarks

Research on AI and economic growth has mostly asked where an economy ends up: at what degree of automation, or at what substitutability between AI and human researchers, growth explodes or settles at a steady rate (Aghion et al., 2019; Jones, 2025; Jones and Tonetti, 2026; Trammell and Korinek, 2020). Some recent contributions do simulate multi-stage adjustment paths (Jones and Tonetti, 2026; Korinek and Suh, 2024; Almeida et al., 2024; Erdil et al., 2025), but less is known analytically about when acceleration starts, what drives it at each stage, when it ends, and what ends it. This paper adds to this literature by developing a semi-endogenous growth model of the multi-stage transitional dynamics of AI-driven automation in scientific research, in which the automation frontier expands endogenously with knowledge, AI capital is accumulated from output, and a physical limit on energy use ends the acceleration.

The key findings were as follows. First, under strong task complementarity ($\sigma < 1$), what AI can do in principle is not the same as how much it actually helps research: as long as some research tasks still need humans, those tasks hold everything back. Second, under the same condition, the effective AI elasticity does not rise gradually. It stays close to zero for decades and then rises sharply once the remaining human bottleneck is passed; the jump takes about ten years in the weak fishing-out scenarios ($\beta = 0.2$), but unfolds over roughly 25–30 years in the baseline with $\beta = 3.1$ and is delayed until 2258 under strong duplication. Third, again under complementarity, the engine of growth changes over time. In the early stage, growth speeds up mainly because tasks are handed from humans to machines; only after the human bottleneck is overcome does building more AI capital become the main driver of faster growth in research effort, and through it of faster idea production and economic growth. Fourth, the outcome depends less on how much work is automated than on how

easily AI can replace human researchers and on how hard it is to find new ideas. Depending on these, the same economy can see no acceleration, a modest transitory acceleration, or an explosive surge.

The paper's limitations need to be acknowledged. It is deliberately aggregate: AI is likely to help some scientific fields far more than others, and AI capital combines very different assets. The automation frontier is calibrated to data on private business as a whole, not to research tasks, and total factor productivity is an imperfect measure of knowledge. In the model AI affects output only through ideas. The automation frontier depends on the level of knowledge, so that interior asymptotic elasticities are attainable only as quasi-steady-state approximations (or under a log-knowledge frontier). The initial automation share is not research-specific, and the investment surplus after the ceiling binds is treated as stranded. Furthermore, the energy limit is a single ceiling, whereas in practice prices and new technologies may ease constraints before any absolute limit is reached.

These limitations suggest items for a future research agenda. In particular, better modelling the economic growth impact of AI will benefit from evidence on how easily AI substitutes for human researchers in research and innovation itself. Future elaborations of the model should allow AI to affect production directly, let the automation frontier respond to investment or depend on proportional changes in knowledge, and allow price-driven extensions of electricity capacity. Finally, the physical limits—thermodynamically—need closer and more detailed modelling.

References

Acemoglu, D. (2025). The simple macroeconomics of AI. *Economic Policy*, 40(121):13–58.

Acemoglu, D. and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30.

Aghion, P., Jones, B. F., and Jones, C. I. (2019). Artificial intelligence and economic growth. In Agrawal, A., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence: An Agenda*, pages 237–290. University of Chicago Press.

Agrawal, A., Gans, J., and Goldfarb, A. (2019). The Economics of Artificial Intelligence: An Agenda. *Chicago: University of Chicago Press*.

Agrawal, A., McHale, J., and Oettl, A. (2023). Superhuman science: How artificial intelligence may impact innovation. *Journal of Evolutionary Economics*, 33:1473–1517.

Almeida, D., Naudé, W., and Sequeira, T. N. (2024). Artificial intelligence and the discovery of new ideas: Is an economic growth explosion imminent? IZA Discussion Paper 16766, IZA Institute of Labor Economics.

Ball, P. (2023). Is AI leading to a reproducibility crisis in science? *Nature*, 624(7990):22–25.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144.

Carpinelli, L., Natoli, F., and Taboga, M. (2026). Artificial Intelligence and the US Economy: An Accounting Perspective on Investment and Production. *Working paper, Bank of Italy*.

Davidson, T. (2021). Report on whether advanced AI could drive explosive economic growth. Technical report, Coefficient Giving.

Davidson, T., Halperin, B., Houlden, T., and Korinek, A. (2026). When Does Automating AI Research Produce Explosive Growth? Feedback Loops in Innovation Networks. *NBER Working Paper 35155*.

Erdil, E. and Besiroglu, T. (2023). Explosive growth from AI automation: A review of the arguments. arXiv preprint 2309.11690, arXiv.

Erdil, E., Potlogea, A., Besiroglu, T., Roldan, E., Ho, A., Sevilla, J., Barnett, M., Vrzla, M., and Sandler, R. (2025). GATE: An Integrated Assessment Model for AI Automation. *arXiv preprint arXiv:2503.04941*.

Gries, T. and Naudé, W. (2022). Modelling Artificial Intelligence in Economics. *Journal for Labour Market Research*, 56(12).

Ho, A., Erdil, E., and Besiroglu, T. (2023). Limits to the Energy Efficiency of CMOS Microprocessors. *In 2023 IEEE International Conference on Rebooting Computing (ICRC) (pp. 1-10). IEEE.*

IEEE (2024). International Roadmap for Devices and Systems: 2024 More Moore. *Institute of Electrical and Electronics Engineers.*

International Energy Agency (2025). Energy and AI. Technical report, International Energy Agency.

Jones, B. (2025). Artificial Intelligence in Research and Development. *NBER Working Paper 34312.*

Jones, C. I. (1995). R&D-based models of economic growth. *Journal of Political Economy*, 103(4):759–784.

Jones, C. I. (2019). Paul Romer: Ideas, nonrivalry, and endogenous growth. *Scandinavian Journal of Economics*, 121(3):859–883.

Jones, C. I. (2022a). The end of economic growth? Unintended consequences of a declining population. *American Economic Review*, 112(11):3489–3527.

Jones, C. I. (2022b). The past and future of economic growth: A semi-endogenous perspective. *Annual Review of Economics*, 14:125–152.

Jones, C. I. and Tonetti, C. (2026). Past automation and future A.I.: How weak links tame the growth explosion. Working paper (Version 0.5), Stanford Graduate School of Business.

Klüppel, L. M. and Knott, A. M. (2023). Are ideas being fished out? *Research Policy*, 52(2):104665.

Korinek, A. and McKelvey, P. (2026). Measuring the AI Economy. *Working Paper 26-9, Peterson Institute for International Economics.*

Korinek, A. and Suh, D. (2024). Scenarios for the transition to AGI. NBER Working Paper 32255, National Bureau of Economic Research.

Naudé, W., Gries, T., and Dimitri, N. (2024). Artificial Intelligence : Economic Perspectives and Models. *Cambridge: Cambridge University Press.*

Nordhaus, W. D. (2021). Are we approaching an economic singularity? Information technology and the future of economic growth. *American Economic Journal: Macroeconomics*, 13(1):299–332.

Rahman, R. (2024). Leading ML Hardware Becomes 40 percent More Energy Efficient Each Year. *Epoch AI Data Insight*.

Ribeiro, B., Meckin, R., Balmer, A., and Shapira, P. (2023). The digitalisation paradox of everyday scientific labour: How mundane knowledge work is amplified and diversified in the biosciences. *Research Policy*, 52(1):104607.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2):S71–S102.

Trammell, P. and Korinek, A. (2020). Economic growth under transformative AI: A guide to the vast range of possibilities for output growth, wages, and the labor share. GPI Working Paper 8-2020, Global Priorities Institute.

Wachter, J. and Wachter, J. (2026). What Investment Data Imply about the AI Transition. *NBER Working Paper 35290*.

Zeira, J. (1998). Workers, machines, and economic growth. *Quarterly Journal of Economics*, 113(4):1091–1117.

Acknowledgment

Tiago Neves Sequeira acknowledges funding from national funds through AI² – Agência para a Investigação e a Inovação, through project UIDB/05037/2025.

Generative AI Disclosure Statement

The authors used Claude Sonnet 5.5 and Gemini 3.8 Flash in the drafting phase to assist with language polishing, grammatical corrections, and stylistic improvements. Claude Sonnet 5.5 was also used to assist with restructuring earlier working papers into this paper, and to help generate the Python script for the simulations, figures and tables. All generated text, code and numerical results were reviewed and edited by the authors, who take full responsibility for the manuscript.

Appendix A: Proofs

Proof of Proposition 1. On a BGP, $g_C = g_K = g_Y$ and log-differentiating (1) gives $g_Y = g_A/(1-\kappa) + n$. With $\varepsilon_{C,t} \to \bar{\varepsilon}_C$ and $\Gamma_t \dot{\alpha}_t \to 0$, (8) gives $g_X = \bar{\varepsilon}_C g_Y + (1-\bar{\varepsilon}_C)n = n + \bar{\varepsilon}_C g_A/(1-\kappa)$. A constant $g_A > 0$ requires $\lambda g_X = \beta g_A$, so $[\beta - \lambda\bar{\varepsilon}_C/(1-\kappa)]g_A = \lambda n$. A positive solution exists if and only if the bracket is positive, which is (12). For stability, holding $\bar{\varepsilon}_C$ fixed, (4) becomes $\dot{g}_A = g_A[\lambda n - (\beta - \lambda\bar{\varepsilon}_C/(1-\kappa))g_A]$, a logistic equation with derivative $-\lambda n < 0$ at the positive steady state. This reduction is evaluated on the balanced manifold, where K_t/Y_t and C_t/Y_t have already converged. Off the BGP, $g_{Y,t} = g_{A,t} + \kappa g_{K,t} + (1-\kappa)n$ and $g_{K,t} = s_K Y_t/K_t - \delta_K$, $g_{C,t} = s_C Y_t/C_t - \delta_C$ evolve with the state variables $(Y_t/K_t, Y_t/C_t)$; the same condition (12) delivers local stability of the coupled system in $(g_{A,t}, Y_t/K_t, Y_t/C_t)$, whose Jacobian we do not reproduce here. ■

Proof of Proposition 2. Let $Q_t \equiv [(1-\alpha_t)/\alpha_t]^{1-\rho}(H_t/C_t)^{\rho}$ so that $\varepsilon_{C,t} = 1/(1+Q_t)$ and $\dot{\varepsilon}_C = -\varepsilon_C(1-\varepsilon_C)\, d\ln Q/dt$. Since $d\ln[(1-\alpha)/\alpha]/dt = -\dot{\alpha}/[\alpha(1-\alpha)] = -\chi\dot{A}/\alpha$ by (6), and $g_H = n$, $d\ln Q/dt = -(1-\rho)\chi A g_A/\alpha + \rho(n - g_C)$. Substituting, and writing $\rho(g_C - n) = -(-\rho)(g_C - n)$ for $\rho < 0$, gives (13). For the limits, $\ln[(1-\alpha_t)/\alpha_t] = \ln(1-\alpha_0) - \chi(A_t - A_0) - \ln\alpha_t$ and $\ln\alpha_t \to 0$ as $A_t \to \infty$, so for $\rho < 0$, $\ln Q_t = \Xi_t + (1-\rho)[\ln(1-\alpha_0) + \chi A_0] + o(1)$. Then $\varepsilon_C = 1/(1+Q) \to 1$ if $\ln Q \to -\infty$, $\to 0$ if $\ln Q \to +\infty$, and to the stated value if $\Xi_t \to \bar{\Xi}$. The condition (14) follows from $\dot{\varepsilon}_C > 0$. ■

Proof of Proposition 3. By (4), $\dot{g}_A > 0$ iff $\lambda g_X > \beta g_A$; substituting (8) gives (15). Before the ceiling binds, $g_{C,t} = s_C Y_t/C_t - \delta_C$ is large when the AI-capital-output ratio is low, which raises the left-hand side; a positive $\Gamma_t\dot{\alpha}_t$ does the same. If capital-output ratios converge, $\varepsilon_{C,t} \to \bar{\varepsilon}_C$ and $\Gamma_t\dot{\alpha}_t \to 0$, the left-hand side converges to βg_A^P and acceleration vanishes. ■

Proof of Proposition 4. The state variables and g_X are continuous, so $\mathcal{S}_t$ is continuous. If $\mathcal{S}_{t_0} = 0$ and $\dot{\mathcal{S}}_{t_0} > 0$, there is $\epsilon > 0$ with $\mathcal{S}_t > 0$ on $(t_0, t_0+\epsilon)$, and $g_A > 0$ implies $\dot{g}_A > 0$ there. If $\mathcal{S}_t$ stays positive on (t_0, T_E), the condition in (16) holds at t_0; it fails for every $t < t_0$ because $\mathcal{S}_{t_0} = 0$ and $t_0 \in (t, T_E)$. Hence $t_0 = T_{\text{TAI}}$. ■

Proof of Proposition 5. Before the ceiling binds, $g_{C,s} = g^u_{C,s}$ and $E_t = E_0 \exp\int_0^t (g^u_{C,s} - d_s)ds$, so $E_t = \bar{E}$ exactly when the integral equals $\ln(\bar{E}/E_0)$; the stated condition guarantees this at a finite date, and if $g^u_C - d \geq \eta$ from t_0 onward, $T_E \leq t_0 + \eta^{-1}\ln(\bar{E}/E_{t_0})$. The ceiling stays binding if and only if the investment rule keeps delivering at least the growth rate d_t required to follow the boundary $C_t = \bar{E}/e_t$, that is $s_C Y_t/(\bar{E}/e_t) - \delta_C \geq d_t$. After T_E, $g_C = d_t \to d$ and (8) gives $g_X \to \bar{\varepsilon}_C d + (1-\bar{\varepsilon}_C)n$. Then (4) is asymptotically $\dot{g}_A = g_A\{\lambda[\bar{\varepsilon}_C d + (1-\bar{\varepsilon}_C)n] - \beta g_A\}$, whose positive stationary point, when the bracketed drift is

positive, is globally stable on $g_A > 0$, which gives (17); g_Y follows from (1) given $g_{K,t} \to g_Y$, which holds because $\dot{K}_t = s_K Y_t - \delta_K K_t$ implies convergence of K_t/Y_t. With $d_t \to 0$ the stationary value is $\lambda(1-\bar{\varepsilon}_C)n/\beta$. When $\bar{\varepsilon}_C = 1$ and $d = 0$ the drift is zero and the equation reduces to $\dot{g}_A = -\beta g_A^2$, which has no positive stationary point: g_A converges globally on $g_A > 0$ to the boundary point 0 with $g_{A,t} \sim (\beta t)^{-1}$, and $A_t \sim \beta^{-1} \ln t \to \infty$. ■

Proof of Proposition 6. The result follows from the definitions of T_{TAI} and T_E and continuity of the state variables. Before T_E, AI capital follows the accumulation law; at T_E its level is continuous and its feasible growth rate switches to d_t. The sign of $\mathcal{S}_t$ determines whether knowledge growth accelerates on either side. If neither boundary is crossed the partial-AI dynamics remain in force (Proposition 1); if only the acceleration boundary is crossed, the economy remains in TAI; if the energy boundary is crossed first, or $\mathcal{S}_t$ returns to non-positive values before T_E, then $T_{\text{TAI}} = \infty$ under (16) and no persistent TAI interval occurs; post-ceiling convergence follows from Proposition 5. ■

Appendix B: Python Code for Replication

The code below reproduces the simulations, figures and tables in this paper. It requires Python 3.9+ with `numpy`, `scipy` and `matplotlib`. To replicate, save the listings as the file names shown, run `python3 figs1.py` and then `python3 tables1.py` from one folder with a subfolder `out/`; all outputs are written to `out/`. All simulations use RK4 with `dt=0.02`. `model2.py` is the simulator used for all figures and tables; its optional extensions (`F`, `cap`) default to `None` and are not used for the results in the paper, so that it reduces to the baseline of `model.py`. The reallocation term is computed from its closed form. `tables1.py` also contains the estimation of $\widehat{\chi}$ behind Table 1.

`model.py`: Baseline simulator

```
"""Simulation of the  model (annual horizon, RK4).
Reproduces the four scenarios of the paper."""
import numpy as np

P0 = dict(n=0.01, kappa=0.35, sK=0.20, dK=0.05, KY0=2.5, A0=2.989, gA0=0.01,
          alpha0=0.75, chi=0.565, rho=-4.0, epsC0=0.10, CY0=0.01, dC=0.20,
          E0frac=0.25, d0=0.15, dinf=0.02, mud=0.04)

SCEN = {
 'Bloom et al.':       dict(lam=1.0, beta=3.1, gC0=0.5),
 'Duplication':        dict(lam=0.2, beta=3.1, gC0=0.5),
 "Moore's Law":        dict(lam=1.0, beta=0.2, gC0=0.5),
 'AI Boom':            dict(lam=1.0, beta=0.2, gC0=1.4),
}

def dpath(t, p):
    return p['dinf'] + (p['d0']-p['dinf'])*np.exp(-p['mud']*t)

def eint(t, p):  # integral of d_s ds
    return p['dinf']*t + (p['d0']-p['dinf'])*(1-np.exp(-p['mud']*t))/p['mud']

def setup(sc, **over):
    p = dict(P0); p.update(SCEN[sc]); p.update(over)
    a0, rho = p['alpha0'], p['rho']
    C0 = p['CY0']*1.0
    # H0 from epsC0
    r = (1-p['epsC0'])/p['epsC0']
    # (1-a)^(1-rho) H^rho = r a^(1-rho) C^rho
    H0 = ((r*a0**(1-rho)*C0**rho)/((1-a0)**(1-rho)))**(1/rho)
    X0 = (a0**(1-rho)*C0**rho/p['epsC0'])**(1/rho)
    p['C0'], p['H0'], p['X0'] = C0, H0, X0
    p['theta'] = p['gA0']*p['A0']**p['beta']/X0**p['lam']
    p['K0'] = p['KY0']
    p['LY0'] = (1.0/(p['A0']*p['K0']**p['kappa']))**(1/(1-p['kappa']))
    p['sC'] = (p['gC0']+p['dC'])*C0/1.0
    p['Ebar'] = p['E0frac']**-1 * C0  # e0=1
    return p

def alpha_of(A, p):
    return 1-(1-p['alpha0'])*np.exp(-p['chi']*(A-p['A0']))

def X_of(C, H, a, p):
    rho = p['rho']
    return (a**(1-rho)*C**rho+(1-a)**(1-rho)*H**rho)**(1/rho)

```

```
def outputs(A, K, C, t, p):
    H = p['H0']*np.exp(p['n']*t); LY = p['LY0']*np.exp(p['n']*t)
    a = alpha_of(A, p)
    X = X_of(C, H, a, p)
    Y = A*K**p['kappa']*LY**(1-p['kappa'])
    gA = p['theta']*X**p['lam']*A**(-p['beta'])
    eps = a**(1-p['rho'])*C**p['rho']/(a**(1-p['rho'])*C**p['rho']+(1-a)**(1-p['rho'])*H**p['rho'])
    return dict(H=H, a=a, X=X, Y=Y, gA=gA, eps=eps)

def simulate(sc, T=300, dt=0.02, **over):
    """Baseline simulator: fixed energy ceiling Ebar, exogenous efficiency path d_t."""
    p = setup(sc, **over)
    n_steps = int(round(T/dt))
    A, K, C = p['A0'], p['K0'], p['C0']
    rec = {k: [] for k in ['t','A','K','C','Y','gA','gY','gC','gK','gX','eps','a','E','bind','S','accum','
     realloc']}
    bind = False; TE = None
    def rhs(state, t):
        A, K, C = state
        o = outputs(A, K, C, t, p)
        dA = o['gA']*A
        dK = p['sK']*o['Y']-p['dK']*K
        dC = p['sC']*o['Y']-p['dC']*C
        return np.array([dA, dK, dC]), o
    for i in range(n_steps+1):
        t = i*dt
        e = np.exp(-eint(t, p))
        o = outputs(A, K, C, t, p)
        gK = p['sK']*o['Y']/K-p['dK']
        # growth of C
        if bind:
            gC = dpath(t, p)
        else:
            gC = p['sC']*o['Y']/C-p['dC']
        gY = o['gA'] + p['kappa']*gK + (1-p['kappa'])*p['n']
        rec['t'].append(t); rec['A'].append(A); rec['K'].append(K); rec['C'].append(C)
        rec['Y'].append(o['Y']); rec['gA'].append(o['gA']); rec['gY'].append(gY)
        rec['gC'].append(gC); rec['gK'].append(gK); rec['eps'].append(o['eps']); rec['a'].append(o['a'])
        rec['E'].append(e*C/p['Ebar']); rec['bind'].append(bind)
        # step (RK4 on A,K,C) with C possibly overwritten by envelope
        if i == n_steps: break
        def f(s, tt):
            return rhs(s, tt)[0]
        s = np.array([A, K, C])
        if bind:
            # C follows envelope: C = Ebar/e_t ; integrate A,K with C prescribed
            def f2(s2, tt):
                Cc = p['Ebar']*np.exp(eint(tt, p))
                return rhs(np.array([s2[0], s2[1], Cc]), tt)[0][:2]
            s2 = s[:2]
            k1 = f2(s2, t); k2 = f2(s2+0.5*dt*k1, t+0.5*dt); k3 = f2(s2+0.5*dt*k2, t+0.5*dt); k4 = f2(s2+dt
     *k3, t+dt)
            s2n = s2 + dt/6*(k1+2*k2+2*k3+k4)
            A, K = s2n; C = p['Ebar']*np.exp(eint(t+dt, p))
        else:
            k1 = f(s, t); k2 = f(s+0.5*dt*k1, t+0.5*dt); k3 = f(s+0.5*dt*k2, t+0.5*dt); k4 = f(s+dt*k3, t+
     dt)
            sn = s + dt/6*(k1+2*k2+2*k3+k4)
            A, K, C = sn
            enew = np.exp(-eint(t+dt, p))
            if enew*C >= p['Ebar']:
                # locate crossing by linear interpolation in log E
                E_old = e*s[2]; E_new = enew*C
                frac = (np.log(p['Ebar'])-np.log(E_old))/(np.log(E_new)-np.log(E_old))
                TE = t+frac*dt
                bind = True
                C = p['Ebar']/enew
    for k in rec: rec[k] = np.array(rec[k])
```

```
    rec['TE'] = TE; rec['p'] = p
    # growth of X, decomposition
    t = rec['t']
    H = p['H0']*np.exp(p['n']*t)
    a = rec['a']; X = np.array([X_of(rec['C'][j], H[j], a[j], p) for j in range(len(t))])
    rec['X'] = X
    # exact decomposition of g_X (Eqs. 8-9): reallocation term in closed form,
    # Gamma*alphadot = (1-rho)/rho * (eps-alpha)/alpha * chi * A * gA (no finite-difference residual)
    rho = p['rho']
    rec['accum'] = rec['eps']*rec['gC']+(1-rec['eps'])*p['n']
    rec['realloc'] = (1-rho)/rho*(rec['eps']-a)/a*p['chi']*rec['A']*rec['gA']
    rec['gX'] = rec['accum']+rec['realloc']
    rec['S'] = p['lam']*rec['gX']-p['beta']*rec['gA']
    return rec

if __name__ == '__main__':
    for sc in SCEN:
        r = simulate(sc, T=300, dt=0.02)
        i = np.argmax(r['gA']); j = np.argmax(r['gY'])
        print(f"{sc:14s} peak gA {r['gA'][i]*100:6.2f}% at {2023+r['t'][i]:.1f}; peak gY {r['gY'][j
    ]*100:6.2f}% at {2023+r['t'][j]:.1f}; TE={r['TE']}; gA100={r['gA'][int(100/0.02)]*100:.2f}, gY100={r['
    gY'][int(100/0.02)]*100:.2f}; a_end={r['a'][-1]:.3f}")
```

model2.py: Extended simulator

```
"""Extended simulator used for all figures and tables. The optional extensions F (efficiency floor) and cap
(price-responsive electricity capacity) default to None and are NOT used for the results in the paper;
with the defaults this reduces to the baseline model of model.py.
State: A, K, C (operational AI capital), Hc (effective deliverable electricity capacity available to AI).
Binding: C = min(unconstrained desired C, Hc/e)."""
import numpy as np
from model import setup, alpha_of, X_of, outputs, dpath, eint, SCEN, P0

def efree_gain(t, p):          # 1/e_free : cumulative efficiency gain factor
    return np.exp(eint(t, p))

def gain(t, p, F):
    x = efree_gain(t, p)
    if F is None: return x
    G = lambda z: F*(1-np.exp(-z/F))
    return G(x)/G(1.0)

def simulate2(sc, T=300, dt=0.02, F=None, cap=None, **over):
    """cap: None -> fixed capacity Hc=Ebar. dict(gH=baseline capacity growth, chiH, tau, gmax, minf,
     price_u0)"""
    p = setup(sc, **over)
    N = int(round(T/dt))
    A, K, C = p['A0'], p['K0'], p['C0']
    Hc = p['Ebar']
    lag = None if cap is None else max(int(round(cap.get('tau',0)/dt)), 0)
    keys = ['t','A','K','C','Y','gA','gY','gC','gK','eps','a','u','P','gH','Hc','e','bind','d']
    rec = {k: np.zeros(N+1) for k in keys}
    TE = None; bind = False
    Pl = np.ones(N+1)
    for i in range(N+1):
        t = i*dt
        G = gain(t, p, F); e = 1.0/G
        o = outputs(A, K, C, t, p)
        u = e*C/Hc
        # price index (stylised): doubles at full utilisation
        P = max(1.0, 2.0*u)
        Pl[i] = P
        if cap is None:
            gH = 0.0
```

```
        else:
            Pd = Pl[max(i-lag,0)]
            gH = (cap.get('gH',0.0) + cap.get('chiH',0.0)*max(Pd-1.0,0.0))*max(1.0-Hc/(cap['minf']*p['Ebar'
    ]),0.0)
            gH = min(gH, cap.get('gmax',0.10))
        gK = p['sK']*o['Y']/K-p['dK']
        gUnc = p['sC']*o['Y']/C-p['dC']
        if i < N:
            enext = 1.0/gain(t+dt, p, F)
            Hn = Hc*np.exp(gH*dt)
        # determine growth of C actually realised
        gC = gUnc
        if bind:
            gC = np.log((Hn/enext)/C)/dt if i < N else rec['gC'][i-1]
        gY = o['gA']+p['kappa']*gK+(1-p['kappa'])*p['n']
        d_t = -np.log(enext/e)/dt if i < N else 0
        for k, v in zip(keys, [t,A,K,C,o['Y'],o['gA'],gY,gC,gK,o['eps'],o['a'],u,P,gH,Hc,e,float(bind),d_t
    ]):
            rec[k][i] = v
        if i == N: break
        # integrate
        def rhs_AK(s, tt, Cfun):
            A_, K_ = s
            oo = outputs(A_, K_, Cfun(tt), tt, p)
            return np.array([oo['gA']*A_, p['sK']*oo['Y']-p['dK']*K_])
        if bind:
            Cn = Hn/enext
            Cfun = lambda tt: C*(Cn/C)**((tt-t)/dt)
            s = np.array([A, K])
            k1 = rhs_AK(s,t,Cfun); k2 = rhs_AK(s+0.5*dt*k1,t+0.5*dt,Cfun); k3 = rhs_AK(s+0.5*dt*k2,t+0.5*dt
    ,Cfun); k4 = rhs_AK(s+dt*k3,t+dt,Cfun)
            s = s+dt/6*(k1+2*k2+2*k3+k4)
            A, K = s; C = Cn; Hc = Hn
        else:
            s = np.array([A, K, C])
            def f(s3, tt):
                oo = outputs(s3[0], s3[1], s3[2], tt, p)
                return np.array([oo['gA']*s3[0], p['sK']*oo['Y']-p['dK']*s3[1], p['sC']*oo['Y']-p['dC']*s3
    [2]])
            k1 = f(s,t); k2 = f(s+0.5*dt*k1,t+0.5*dt); k3 = f(s+0.5*dt*k2,t+0.5*dt); k4 = f(s+dt*k3,t+dt)
            sn = s+dt/6*(k1+2*k2+2*k3+k4)
            Cn_cap = Hn/enext
            if sn[2] >= Cn_cap:
                # crossing inside step: interpolate in log(u)
                u_old = u; u_new = sn[2]*enext/Hn
                frac = (0-np.log(u_old))/(np.log(u_new)-np.log(u_old))
                TE = t+frac*dt if TE is None else TE
                bind = True
                A, K = sn[0], sn[1]; C = Cn_cap
            else:
                A, K, C = sn
            Hc = Hn
    rec['TE'] = TE; rec['p'] = p
    t = rec['t']; H = p['H0']*np.exp(p['n']*t)
    X = np.array([X_of(rec['C'][j], H[j], rec['a'][j], p) for j in range(N+1)])
    # closed-form decomposition of g_X (no finite-difference residual at T_E)
    rho = p['rho']; a_ = rec['a']
    rec['accum'] = rec['eps']*rec['gC']+(1-rec['eps'])*p['n']
    rec['realloc'] = (1-rho)/rho*(rec['eps']-a_)/a_*p['chi']*rec['A']*rec['gA']
    rec['X'] = X; rec['gX'] = rec['accum']+rec['realloc']; rec['S'] = p['lam']*rec['gX']-p['beta']*rec['gA'
    ]
    return rec

def summarize(r, thr=0.30):
    t = r['t']; dt = t[1]-t[0]
    gY = r['gY']; j = np.argmax(gY)
    above = (gY >= thr).sum()*dt
    TEv = r['TE']
```

```
    def at(x):
        k = int(round(x/dt)); k = min(k, len(t)-1); return gY[k]
    out = dict(TE=TEv, peakgY=gY[j], peakyr=t[j], peakgA=r['gA'].max(), years_above=above)
    if TEv is not None:
        for h in [10, 25, 50, 100, 200]:
            out[f'gY+{h}'] = at(TEv+h) if TEv+h <= t[-1] else np.nan
    out['gY_end'] = gY[-1]; out['gA_end'] = r['gA'][-1]
    return out

if __name__ == '__main__':
    from model import simulate
    for sc in ["Moore's Law","AI Boom"]:
        a = simulate(sc, dt=0.02); b = simulate2(sc, dt=0.02)
        print(sc, 'peak gY', a['gY'].max(), b['gY'].max(), 'TE', a['TE'], b['TE'], 'max|dgY|', np.abs(a['gY
    ']-b['gY']).max())
```

`style.py`: Plot style

```
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
plt.rcParams.update({'font.size':9,'axes.titlesize':9,'axes.labelsize':9,'legend.fontsize':7.5,'xtick.
    labelsize':8,'ytick.labelsize':8,
 'axes.spines.top':False,'axes.spines.right':False,'axes.grid':True,'grid.alpha':0.25,'grid.linewidth':0.5,
    'figure.dpi':150,'savefig.bbox':'tight','font.family':'serif'})
COL={'Bloom et al.':'#1b4f72','Duplication':'#7f8c8d','Moore\'s Law':'#b03a2e','AI Boom':'#d4ac0d'}
LS={'Bloom et al.':'-','Duplication':'--','Moore\'s Law':'-','AI Boom':'-.'}
```

`figs1.py`: Frontier calibration, scenarios, figures and regime map

```
from style import *
from model2 import *
import numpy as np
OUT='./out/'
# ---------- Fig 1: calibrated frontier
al=np.array([0.28,0.45,0.63,0.75]); yrs=np.array([1950,1975,2000,2023]); chi=0.565
fig,ax=plt.subplots(1,2,figsize=(7.2,2.9))
t=np.arange(1950,2024); A=np.exp(0.015*(t-1950)); fit=1-0.72*np.exp(-chi*(A-1))
ax[0].plot(t,fit,'-',color='#1b4f72',label=r'Fitted frontier, $\hat\chi=0.565$'); ax[0].plot(yrs,al,'o',
    color='#b03a2e',label='Jones--Tonetti automation cutoffs')
ax[0].set_ylim(0.2,0.85); ax[0].set_xlabel('Year'); ax[0].set_ylabel(r'Automation share $\alpha_t$'); ax
    [0].legend(loc='lower right'); ax[0].set_title('(a) Fit to 1950--2023 moments')
A23=np.exp(0.015*73); tt=np.arange(2023,2201)
for g,ls,l in [(0.005,':','0.5%'),(0.010,'--','1.0%'),(0.015,'-','1.5%')]:
    Af=A23*np.exp(g*(tt-2023)); ax[1].plot(tt,1-0.72*np.exp(-chi*(Af-1)),ls,color='#1b4f72',label=f'
    Knowledge growth {l}')
for y in [0.9,0.99]: ax[1].axhline(y,color='k',lw=0.5)
ax[1].set_ylim(0.7,1.005); ax[1].set_xlabel('Year'); ax[1].set_ylabel(r'$\alpha_t$'); ax[1].legend(loc='
    lower right'); ax[1].set_title('(b) Projection of the frontier')
fig.tight_layout(); fig.savefig(OUT+'fig1_frontier.pdf'); plt.close()
# ---------- simulate scenarios
R={sc:simulate2(sc,dt=0.02) for sc in SCEN}
yr=lambda r: 2023+r['t']
H=100; sl=lambda r: r['t']<=H
fig,ax=plt.subplots(2,2,figsize=(7.2,5.6))
# (a) knowledge growth, dual axis
a=ax[0,0]; a2=a.twinx(); a2.spines['right'].set_visible(True); a2.grid(False)
for sc in ["Moore's Law","AI Boom"]: a.plot(yr(R[sc])[sl(R[sc])],100*R[sc]['gA'][sl(R[sc])],LS[sc],color=
    COL[sc],label=sc)
for sc in ['Bloom et al.','Duplication']: a2.plot(yr(R[sc])[sl(R[sc])],100*R[sc]['gA'][sl(R[sc])],LS[sc],
    color=COL[sc],label=sc)
```

```
a.set_ylabel(r'Explosive scenarios, $g_A$ (%)'); a2.set_ylabel(r'Aggregate scenarios, $g_A$ (%)'); a.set_title('(a) Knowledge growth')
h1,l1=a.get_legend_handles_labels(); h2,l2=a2.get_legend_handles_labels(); a.legend(h1+h2,l1+l2,loc='upper left',fontsize=6.5)
a2.set_ylim(0,2.2)
# (b) output growth
a=ax[0,1]; a2=a.twinx(); a2.spines['right'].set_visible(True); a2.grid(False)
for sc in ["Moore's Law","AI Boom"]: a.plot(yr(R[sc])[sl(R[sc])],100*R[sc]['gY'][sl(R[sc])],LS[sc],color=COL[sc])
for sc in ['Bloom et al.','Duplication']: a2.plot(yr(R[sc])[sl(R[sc])],100*R[sc]['gY'][sl(R[sc])],LS[sc],color=COL[sc])
a.set_ylabel(r'Explosive scenarios, $g_Y$ (%)'); a2.set_ylabel(r'Aggregate scenarios, $g_Y$ (%)'); a.set_title('(b) Output growth'); a2.set_ylim(0,4.5)
# (c) energy
a=ax[1,0]
for sc in SCEN: a.plot(yr(R[sc])[sl(R[sc])],R[sc]['u'][sl(R[sc])],LS[sc],color=COL[sc])
a.axhline(1,color='k',lw=0.6); a.set_ylabel(r'$E_t/\bar E$'); a.set_xlabel('Year'); a.set_title('(c) Energy use relative to the ceiling'); a.set_ylim(0,1.1)
# (d) frontier
a=ax[1,1]
for sc in SCEN: a.plot(yr(R[sc])[sl(R[sc])],R[sc]['a'][sl(R[sc])],LS[sc],color=COL[sc])
a.set_ylabel(r'Automation frontier $\alpha_t$'); a.set_xlabel('Year'); a.set_title('(d) Automation frontier'); a.set_ylim(0.7,1.01)
for x in ax[1]: x.set_xlabel('Year')
fig.tight_layout(); fig.savefig(OUT+'fig2_scenarios.pdf'); plt.close()
# ---------- Fig 3: anatomy
fig,ax=plt.subplots(2,2,figsize=(7.2,5.4))
for j,sc in enumerate(['Bloom et al.',"Moore's Law"]):
    r=R[sc]; x=yr(r); m=r['t']<=(100 if j==0 else 60); xx=x[m]
    a=ax[0,j]; a.plot(xx,r['a'][m],'-',color='#1b4f72',label=r'automation frontier $\alpha_t$'); a.plot(xx,r['eps'][m],'--',color='#b03a2e',label=r'effective AI elasticity $\varepsilon_{C,t}$')
    a.axhline(0.5,color='k',lw=0.5)
    ib=np.argmax(r['eps']>=0.5); a.axvline(2023+r['t'][ib],color='k',lw=0.6,ls=':'); a.text(2023+r['t'][ib]+1,0.08,r'$T_B$',fontsize=8)
    if r['TE']: a.axvline(2023+r['TE'],color='#d4ac0d',lw=1.2,ls='-'); a.text(2023+r['TE']+0.5,0.3,r'$T_E$',fontsize=8)
    a.set_title(f"({'ab'[j]}) {sc}: frontier and effective share"); a.set_ylim(-0.02,1.05); a.legend(loc='center left',fontsize=6.5)
    b=ax[1,j]
    n=r['p']['n']; acc=r['eps']*r['gC']; hum=(1-r['eps'])*n; rea=r['realloc']
    lam=r['p']['lam']; be=r['p']['beta']
    comp=[lam*acc[m],lam*hum[m],np.maximum(lam*rea[m],0)]
    b.stackplot(xx,*[100*c for c in comp],labels=[r'AI accumulation $\lambda\varepsilon_{C,t} g_C$',r'human research $\lambda(1-\varepsilon_{C,t})n$',r'task reallocation $\lambda\Gamma_t\dot\alpha_t$ (if $>0$)'],colors=['#5dade2','#aab7b8','#f5b041'],alpha=0.85)
    b.plot(xx,100*be*r['gA'][m],'k-',lw=1.2,label=r'fishing-out drag $\beta g_A$')
    b.set_title(f"({'cd'[j]}) {sc}: terms of $\\lambda g_X$ vs $\\beta g_A$"); b.set_ylabel('% per year'); b.set_xlabel('Year')
    if r['TE']: b.axvline(2023+r['TE'],color='#d4ac0d',lw=1.2)
    b.legend(loc='upper left',fontsize=6.2)
fig.tight_layout(); fig.savefig(OUT+'fig3_anatomy.pdf'); fig.savefig(OUT+'Figure3_anatomy.png',dpi=200); plt.close()
# ---------- Fig 4: regime map (sigma x beta)
sig=np.array([0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8]); betas=np.array([0.2,0.5,1.0,2.0,3.1])
Z=np.zeros((len(betas),len(sig))); TEm=np.full((len(betas),len(sig)),np.nan)
for i,b in enumerate(betas):
    for k,s in enumerate(sig):
        r=simulate2("Moore's Law",dt=0.02,beta=b,rho=1-1/s); sm=summarize(r)
        Z[i,k]=sm['peakgY']; TEm[i,k]=sm['TE'] if sm['TE'] else np.nan
np.save('./out/regimemap_Z.npy',Z); np.save('./out/regimemap_TE.npy',TEm)
fig,ax=plt.subplots(figsize=(5.6,3.3))
im=ax.imshow(np.log10(100*Z),origin='lower',aspect='auto',cmap='YlOrRd',extent=[0,len(sig),0,len(betas)],vmin=0,vmax=4)
for i in range(len(betas)):
    for k in range(len(sig)):
        v=100*Z[i,k]; lab=(f"{v:.1f}" if v<10 else (f"{v:.0f}" if v<1e4 else r"$>10^4$")); ax.text(k+0.5,i+0.5,lab+('*' if not np.isnan(TEm[i,k]) else ''),ha='center',va='center',fontsize=6.5)
ax.set_xticks(np.arange(len(sig))+0.5); ax.set_xticklabels([f"{s:.1f}" for s in sig]); ax.set_yticks(np.
```

```
    arange(len(betas))+0.5); ax.set_yticklabels([f"{b:g}" for b in betas])
ax.set_xlabel(r'Elasticity of substitution $\sigma=1/(1-\rho)$'); ax.set_ylabel(r'Fishing-out parameter $\
    beta$'); ax.grid(False)
cb=fig.colorbar(im,ax=ax); cb.set_label(r'$\log_{10}$ peak output growth (%)')
fig.tight_layout(); fig.savefig(OUT+'fig4_regimemap.pdf'); plt.close()
print('done')
```

tables1.py: Results and robustness tables

```
from model2 import *
import numpy as np, json, math
OUT='./out/'
rh=lambda x: int(math.floor(x+0.5))
def tb(r):
    e=r['eps']; return r['t'][np.argmax(e>=0.5)] if (e>=0.5).any() else None
R={sc:simulate2(sc,dt=0.02) for sc in SCEN}
rows=[]; info={}
for sc,r in R.items():
    i=np.argmax(r['gA']); j=np.argmax(r['gY']); TB=tb(r); TE=r['TE']
    g100=r['gY'][int(100/0.02)]; gA100=r['gA'][int(100/0.02)]
    gTE100=r['gY'][int(min(TE+100,299.9)/0.02)] if TE else None
    a95=r['t'][np.argmax(r['a']>=0.95)]
    info[sc]=dict(gA_pk=r['gA'][i],gA_yr=2023+r['t'][i],gY_pk=r['gY'][j],gY_yr=2023+r['t'][j],TB=TB,TE=TE,
     g100=g100,gTE100=gTE100,a95=a95,gY300=r['gY'][-1],gA300=r['gA'][-1],gY200=r['gY'][int(200/0.02)])
    rows.append(f"{sc} & {100*r['gA'][i]:.1f}\\% ({rh(2023+r['t'][i])}) & {100*r['gY'][j]:.1f}\\% ({rh
     (2023+r['t'][j])}) & "+(f"{rh(2023+TB)}" if TB else "---")+" & "+(f"{rh(2023+TE)}" if TE else "Does
     not bind")+" & "+f"{100*g100:.1f}\\% \\\\")
open(OUT+'tab_results.tex','w').write('\n'.join(rows)+'\n')
json.dump({k:{a:(float(b) if b is not None else None) for a,b in v.items()} for k,v in info.items()},open('
    ./out/res_p1_info.json','w'),indent=1)
# Table 1: calibration of chi to the four automation moments for alternative growth of knowledge
from scipy.optimize import minimize_scalar
al=np.array([0.28,0.45,0.63,0.75]); yrs=np.array([1950,1975,2000,2023]); rows=[]
for g in [0.010,0.012,0.015,0.020]:
    A=np.exp(g*(yrs-1950)); sse=lambda c: np.sum((al-(1-0.72*np.exp(-c*(A-1))))**2)
    c=minimize_scalar(sse,bounds=(0.01,5),method='bounded',options={'xatol':1e-10}).x
    rm=np.sqrt(sse(c)/4)
    # year alpha=0.90 if knowledge keeps growing at g from the fitted 2023 share
    tt=np.arange(2023,2400,0.01); Af=A[-1]*np.exp(g*(tt-2023)); y90=tt[np.argmax(1-0.72*np.exp(-c*(Af-1))
     >=0.90)]
    rows.append((g,c,rm,y90))
open(OUT+'tab_chi.txt','w').write('\n'.join(f"g={g:.3f} chi={c:.3f} rmse={rm:.3f} y90={y90:.0f}" for g,c,rm
    ,y90 in rows)+'\n')
print(open(OUT+'tab_chi.txt').read())
# robustness grid
rows=[]
for chi in [0.357,0.565,0.995]:
    for sig in [0.1,0.2,0.5]:
        r=simulate2("Moore's Law",dt=0.02,chi=chi,rho=1-1/sig); s=summarize(r); TB=tb(r)
        r2=simulate2("Bloom et al.",dt=0.02,chi=chi,rho=1-1/sig); s2=summarize(r2); TB2=tb(r2)
        rows.append(f"{chi:.3f} & {sig:.1f} & {TB:.1f} & {s['TE']:.1f} & {100*s['peakgY']:.0f}\\% & {TB2:.1
    f} & {100*s2['peakgY']:.1f}\\% \\\\")
open(OUT+'tab_robust.tex','w').write('\n'.join(rows)+'\n')
print(open(OUT+'tab_results.tex').read()); print(open(OUT+'tab_robust.tex').read())
for k,v in info.items(): print(k,{a:(round(b,3) if b is not None else None) for a,b in v.items()})
```