

The Upside of Being Replaceable: AI, Market Structure, and Growth

Oliver Pardo

Departamento de Administración

Pontificia Universidad Javeriana

pardoo@javeriana.edu.co

September 2026

Abstract

I develop an endogenous-growth model in which a monopolistic AI developer supplies AI services and invests in recursive AI self-improvement. I characterize two long-run regimes under an upper bound on AI efficiency—the quantity of AI services produced per unit of compute. Complementarity between AI and labor, or a sufficiently low upper bound on AI efficiency, can leave labor as a production bottleneck. In this regime, output per worker and wages eventually grow at the exogenous rate of labor-augmenting technological progress. Instead, when AI and labor are substitutes and the upper bound is sufficiently high, an AI-dominated regime sustains permanently faster wage growth and a higher long-run return to capital. Output per worker grows faster than wages, so labor’s income share vanishes while AI-industry revenue increases as a share of output. Replacing monopoly with perfect competition eliminates private research incentives. Nevertheless, if AI is already sufficiently efficient, the competitive economy can sustain AI-dominated growth through capital accumulation and expanding AI use. By contrast, monopoly rewards research but restricts AI use through its markup. It may therefore produce more efficient AI yet slower long-run growth than competition. Numerical simulations show that early output-per-worker growth can remain moderate and similar across economies approaching different long-run regimes.

Keywords: Artificial Intelligence; Recursive AI self-improvement; Market Structure; Labor Share.

JEL codes: O41, O33, O31, D42, E25.

[Digital edition and simulations](#) | [Latest PDF version](#)

Contents

1	Introduction	3
2	Related literature	5
3	Model	6
3.1	Population and labor-augmenting technology	6
3.2	Household	6
3.3	Final-good production	7
3.4	The AI developer	8
3.5	Equilibrium	12
4	Equilibrium with an upper bound on AI efficiency	13
4.1	The return to capital and the efficiency threshold	14
4.2	Equilibrium existence and long-run regimes	15
4.3	Substitution and wage growth	17
5	Competitive Equilibrium	19
5.1	Characterization	19
5.2	Comparison with monopoly	20
5.3	Research incentives and policy	22
6	Numerical simulations	23
6.1	Parameters and initial conditions	23
6.2	Fast takeoff	24
6.3	Slow takeoff	29
6.4	From competition to monopoly	33
7	Conclusion	35
A	Proofs	36
A.1	Developer optimization	36
A.2	Equilibria with bounded AI efficiency	40
A.3	Competitive equilibrium and comparison with monopoly	49
B	Numerical algorithm and equilibrium audit	53
B.1	Numerical solution	53
B.2	Accuracy and equilibrium checks	54
C	Equilibrium without an upper bound on AI efficiency	55
C.1	Complementarity	55
C.2	Unit elasticity	55
C.3	Substitution	57
	References	58

1 Introduction

Advances in artificial intelligence raise both hopes and fears. One hope is that increasingly capable AI could generate unprecedented economic prosperity, with workers benefiting from higher wages. One fear is that the AI industry could capture a growing share of that prosperity while workers who depend on labor income receive a declining share. In this paper, I show how both outcomes can coexist: when AI and labor are substitutes, AI can permanently raise wage growth. Yet in the long run, output per worker grows even faster than wages, while AI-industry revenue grows faster than total labor income. Labor’s income share therefore converges to zero.

To show this, I extend the Ramsey–Cass–Koopmans model to include AI services as an additional production input. Competitive firms produce the final good using capital, effective labor, and AI services. A monopolistic AI developer supplies the AI services and invests in recursive AI self-improvement (RSI) to improve AI efficiency—the quantity of AI services produced per unit of compute. Through RSI, more efficient AI helps develop still more efficient AI. A representative household chooses consumption and saving, receiving income from labor, capital ownership, and the AI developer’s profits.

If labor and AI are substitutes and there are sufficiently strong research returns, the developer’s profit can become unbounded if AI efficiency can grow without bound. To rule this out, I first study an economy with an upper bound on AI efficiency and examine how equilibria change as the bound rises. Together, the elasticity of substitution between AI and labor and the upper bound on AI efficiency determine whether human labor remains a long-run production bottleneck or whether AI sustains permanently faster economic growth.

In the equilibria where labor remains a long-run production bottleneck, output per worker and wages eventually grow at the exogenous rate of labor-augmenting technological progress, and labor’s income share converges to a positive value. Such equilibria can arise under complementarity between labor and AI production services or when the upper bound on AI efficiency is sufficiently low.

An AI-dominated long-run regime becomes possible when AI and labor are substitutes and the upper bound on AI efficiency is sufficiently high. As in [Jones and Manuelli \(1990\)](#), capital accumulation can sustain growth when its limiting return remains sufficiently high. In this regime, capital and AI production services eventually grow faster than effective labor. This does not require AI efficiency to keep improving: capital accumulation and expanding compute use sustain growth even as AI efficiency approaches its upper bound. The long-run growth rates of output per worker and wages, as well as the interest rate, exceed their labor-bottleneck values. Because output per worker grows faster than wages, labor’s income share converges to zero. Total labor income therefore grows faster than in the labor-bottleneck regime, but in the long run it grows more slowly than capital income and AI-industry revenue. Greater substitutability raises long-run output-per-worker growth, but long-run wage growth peaks at a finite elasticity of substitution.

I also characterize a competitive economy in which AI models and any improvements are freely available to all suppliers of AI production services. Without a private return to research, AI efficiency remains at its initial level. When AI and labor are substitutes, a sufficiently high initial efficiency

nevertheless sustains AI-dominated growth through capital accumulation and expanding compute use. Neither monopoly nor continuing RSI is therefore necessary to sustain this growth regime.

The growth comparison between monopoly and competition depends on whether monopoly-financed efficiency gains compensate for the reduction in AI use caused by monopoly pricing. A monopoly economy may reach AI-dominated growth from an initial efficiency too low to sustain it under competition. Conversely, competition can sustain AI-dominated growth from initial states for which monopoly-led research improves efficiency but leaves the economy labor-bottlenecked. Even when both economies approach AI-dominated growth, monopoly delivers faster long-run growth only if the research gains are large enough to offset its markup.

I also examine the effect of removing the upper bound on AI efficiency in Appendix C. With complementarity, long-run average output-per-worker growth cannot exceed the exogenous rate of labor-augmenting technological progress. At unit elasticity, I derive conditions for a balanced-growth equilibrium in which output per worker grows faster than this exogenous rate. Wages grow at the same rate as output per worker, preserving a constant labor income share. With substitution and sufficiently strong research returns, the developer can make discounted net profit arbitrarily large even over a fixed finite horizon, ruling out the existence of an equilibrium.

I use numerical simulations to illustrate the effects of introducing RSI under monopoly and replacing competition with monopoly. First, I simulate an unanticipated activation of RSI in economies that already use AI services under a monopoly. Depending on research productivity, output-per-worker growth can display a moderate increase or a noticeable jump in both long-run regimes. Eventually, output-per-worker growth converges to the rate of labor-augmenting technological progress in the labor-bottleneck regime and to a permanently higher rate in the AI-dominated regime. With lower research productivity, initial output-growth trajectories can remain similar across economies approaching different long-run regimes. Second, I simulate granting exclusive AI rights in economies that initially supply AI services competitively. Monopoly-financed research can move an economy from labor-bottleneck to AI-dominated growth. Conversely, monopoly pricing can leave an economy labor-bottlenecked even when competition would sustain AI-dominated growth with less efficient AI.

This paper has seven sections, including the present introduction. Section 2 relates the paper to the literature. Section 3 presents the model and defines the equilibrium. Section 4 establishes sufficient conditions for equilibrium existence with an upper bound on AI efficiency and characterizes the associated long-run growth and distributional outcomes. Section 5 characterizes competitive AI provision and compares its growth outcomes with those under monopoly. Section 6 presents numerical simulations. Section 7 concludes. Appendices A and B contain the proofs and describe the algorithm used for the simulations. Appendix C examines the three substitution regimes without an upper bound on AI efficiency.

2 Related literature

My model shares the feedback between AI development and economic growth studied by [Davidson et al. \(2026\)](#), [Jones and Tonetti \(2026\)](#), and [Erdil et al. \(2025\)](#): better AI raises output, expanding the resources available for further AI development. My model also allows better AI to improve research itself through RSI, a technological feedback related to the automation of research in [Davidson et al. \(2026\)](#) and [Jones and Tonetti \(2026\)](#). The models differ in how saving and research expenditure are determined. [Davidson et al. \(2026\)](#) and [Jones and Tonetti \(2026\)](#) take saving and research allocation shares as given in their growth models. [Erdil et al. \(2025\)](#) instead let a social planner choose investment and allocate compute between developing and deploying AI. I instead study a market equilibrium.

I examine when AI can sustain faster growth rather than raise it only temporarily. [Aghion et al. \(2019\)](#) and [Jones and Tonetti \(2026\)](#) show how tasks that remain dependent on human labor can constrain growth despite automation elsewhere. [Palivos and Karagiannis \(2010\)](#) show how sufficiently high substitutability between capital and labor can allow capital accumulation to sustain growth without technological progress. [Restrepo \(2025\)](#) obtains sustained output growth from an exogenous expansion of compute after bottleneck tasks have been automated. In my model, RSI can improve AI efficiency without eliminating this labor bottleneck. I also identify conditions under which capital and AI services expand together and sustain permanently faster growth without further AI-efficiency gains.

I also study how AI-driven growth affects wages and labor’s share of income. [Acemoglu and Restrepo \(2018\)](#) show that capital accumulation can raise long-run wages even as automation displaces labor from existing tasks and reduces labor’s income share. Similarly, [Berg et al. \(2018\)](#) show that, when robots and labor are imperfect substitutes, higher robot productivity and capital accumulation can raise steady-state wages while lowering labor’s share. These comparative-static results concern wage levels, whereas I examine when AI can permanently increase wage growth. Like [Ray and Mookherjee \(2022\)](#), I show that wages can rise without bound even as labor’s income share vanishes. [Hémous and Olsen \(2022\)](#) show that low-skill wages grow more slowly than output in the long run, while high-skill wages keep pace with it. In this paper, I abstract from labor heterogeneity.

I also study how market structure affects AI development and growth. [Korinek and Ng \(2019\)](#) examine how digital innovation lowers production costs while generating monopoly rents and markups that restrict output. Holding AI capability fixed, [Gans \(2026\)](#) shows how competition and monopoly can affect long-run growth by changing the choice between incremental and exploratory research. [Liu and Wan \(2026\)](#) show how the prices charged by general-purpose AI suppliers affect incentives to improve both their models and the specialized applications that use them, shaping long-run growth. I compare monopoly and competition to examine whether monopoly-led research raises AI efficiency enough to compensate for the reduction in AI use caused by monopoly pricing.

3 Model

Time is continuous. The economy contains a representative household, competitive final-good firms, and an integrated AI developer. Labor-augmenting technology A grows exogenously. AI efficiency B improves through autonomous AI research: the only current input used to raise B is research compute. The model therefore describes an economy in which there are no longer human researchers developing AI and AI has achieved recursive self-improvement (RSI). As in [Trammell and Korinek \(2023\)](#), I distinguish automation of research from automation of production: human labor remains an input in final-good production.

3.1 Population and labor-augmenting technology

Population grows at the exogenous rate $n \geq 0$:

$$\dot{N} = nN, \quad n \geq 0. \quad (1)$$

Each person supplies one unit of homogeneous labor inelastically. The aggregate labor endowment is therefore N . Because autonomous research uses no human labor, all labor is employed in final production and equilibrium will impose $L = N$.

The variable A measures the level of labor-augmenting technology. One unit of labor supplies A efficiency units, so effective labor is AL . Labor-augmenting technology evolves according to

$$\dot{A} = \gamma A, \quad \gamma > 0. \quad (2)$$

3.2 Household

The representative household owns physical capital and the AI developer. Let Π denote the AI developer's net profit, which accrues to the household. The household takes the paths of the net interest rate r , the wage w in final-good units and Π as given. For a predetermined capital stock $K_0 > 0$, it chooses aggregate consumption C to solve

$$\max_{\{C_t\}_{t \geq 0}} \int_0^\infty e^{-\rho t} N_t \log(C_t/N_t) dt, \quad \rho > n, \quad (3)$$

subject to

$$\dot{K} = rK + wN + \Pi - C, \quad K \geq 0. \quad (4)$$

The variable r is the net real interest rate: depreciation has already been deducted from the rental return to capital. The household's Euler condition is

$$\frac{\dot{C}}{C} = n + r - \rho. \quad (5)$$

Its transversality condition is

$$\lim_{t \rightarrow \infty} e^{-\rho t} \frac{N_t K_t}{C_t} = 0. \quad (6)$$

3.3 Final-good production

Competitive firms produce the final good from physical capital K , effective labor AL , and effective AI production services X . Let $0 < \alpha < 1$ denote the capital share, $\delta \geq 0$ the depreciation rate, and $\sigma > 0$ the elasticity of substitution between effective labor and effective AI production services. The parameters ω_L and ω_X denote the CES weights on effective labor and effective AI production services, respectively. Both weights are positive and sum to one. Production is

$$Y = K^\alpha Z^{1-\alpha}, \quad (7)$$

$$Z = \left[\omega_L (AL)^{(\sigma-1)/\sigma} + \omega_X X^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}. \quad (8)$$

The composite input Z combines effective labor and AI production services.

I use the CES specification in equation (8) to represent how labor and AI services combine, either within a single task or as aggregate inputs across multiple tasks. The latter interpretation connects to [Jones and Tonetti \(2026\)](#), who show how a task-based production process can admit a CES-like aggregate representation.

Let p_X denote the price of one unit of effective AI production services. Taking the input prices $r + \delta$, w , and p_X as given, a final-good firm solves

$$\max_{K, L, X \geq 0} \left\{ K^\alpha Z(AL, X)^{1-\alpha} - (r + \delta)K - wL - p_X X \right\}. \quad (9)$$

Define s_X as the elasticity of Z with respect to X :

$$s_X = \frac{\omega_X X^{(\sigma-1)/\sigma}}{\omega_L (AL)^{(\sigma-1)/\sigma} + \omega_X X^{(\sigma-1)/\sigma}}. \quad (10)$$

The first-order conditions are

$$r = \alpha \frac{Y}{K} - \delta, \quad (11)$$

$$w = (1 - \alpha)(1 - s_X) \frac{Y}{L}, \quad (12)$$

$$p_X = (1 - \alpha)s_X \frac{Y}{X}. \quad (13)$$

Thus $(1 - \alpha)s_X$ is AI revenue as a share of output, while $(1 - \alpha)(1 - s_X)$ is labor's income share.

Constant returns and these conditions imply

$$Y = (r + \delta)K + wL + p_X X. \quad (14)$$

Solving the final-good firm's AI condition for its price yields the inverse demand schedule $p_X(X)$. The absolute elasticity of inverse demand, holding K , A , and L fixed, is

$$e_X \equiv -\frac{\partial \ln p_X(X)}{\partial \ln X} = \frac{1 - s_X}{\sigma} + \alpha s_X. \quad (15)$$

This demand schedule is an input to the developer's problem.

3.4 The AI developer

3.4.1 Effective AI production and research services

I use *compute* to denote the computational resources used by AI systems (Sastry et al., 2024). Compute can be measured in floating-point operations, with flows measured in operations per unit of time. I assume that each unit of compute requires one unit of the final good, which could otherwise be consumed or invested in physical capital.

AI efficiency B measures the services obtained per unit of compute. A rise in B represents algorithmic improvement rather than an expansion of physical computing capacity. This interpretation parallels measures of algorithmic efficiency based on the compute required to achieve a given level of performance (Hernandez and Brown, 2020; Erdil and Besiroglu, 2022).

Inference compute U supplies AI services for current final-good production, while research compute M is used to develop better AI. Korinek and McKelvey (2026) make a related distinction between compute used for inference and compute used for R&D and training. Research compute includes both training and running AI systems that conduct research (Davidson et al., 2026). I assume that the same B augments both activities, producing flows BU and BM of AI production and research services, respectively. AI production services therefore satisfy

$$X = BU. \quad (16)$$

Figure 1 summarizes the production block.

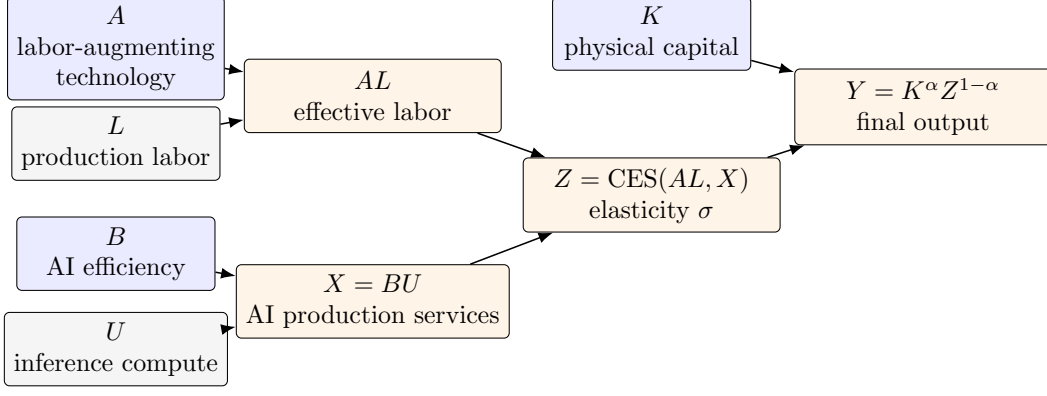


Figure 1: Production of final output and AI production services. AI efficiency B augments inference compute U to produce $X = BU$, the flow of AI production services. These services combine with effective labor AL to produce Z , which combines with capital K to produce Y .

3.4.2 RSI technology

Given initial AI efficiency $B_0 > 0$ and an upper bound $\bar{B} \in (B_0, \infty]$, AI efficiency evolves according to the RSI technology

$$\dot{B} = \chi(BM)^\eta \psi(B), \quad \psi(B) = 1 - \frac{B}{\bar{B}}, \quad \chi > 0, \quad 0 < \eta < 1. \quad (17)$$

Figure 2 summarizes the RSI mechanism.

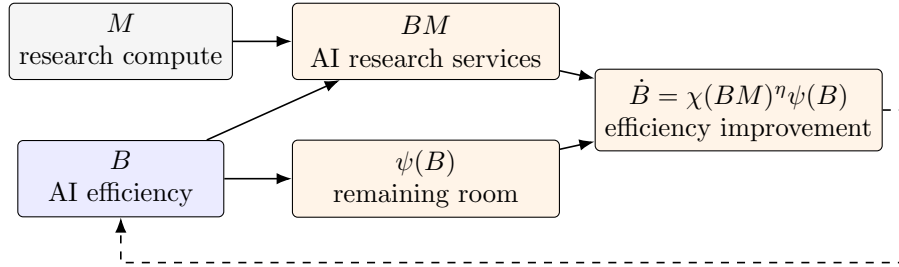


Figure 2: Recursive self-improvement in AI efficiency. AI efficiency B augments research compute M to produce BM , the flow of AI research services. These services and the remaining room for improvement $\psi(B)$ determine $\dot{B}$, which increases B .

The parameter χ shifts research productivity, while η is the elasticity of $\dot{B}$ with respect to research compute, holding B fixed. Equation (17) captures RSI: a higher B produces more AI research services BM from a given amount of research compute M , and these services help improve B further. This is the AI-research feedback studied by Davidson et al. (2026). The term $\psi(B)$ reduces the remaining room for improvement as AI efficiency approaches $\bar{B}$. A related approach appears in Erdil et al. (2025), who also reduce research productivity as hardware and software efficiency approach their respective upper bounds. The uncapped specification is recovered at $\bar{B} = \infty$.

Evidence that ideas become harder to find motivates allowing research productivity to decline as technology advances. [Bloom et al. \(2020\)](#) document that research effort has risen substantially while research productivity has fallen across several industries, products, and firms. This evidence motivates diminishing returns, but it does not establish a literal finite technological ceiling. The term $1 - B/\bar{B}$ is a deliberately stronger reduced-form specification. Without the upper bound, recursive improvement may lead to the absence of an equilibrium, as shown in [Appendix C.3](#).

3.4.3 Supply of effective AI production services

An integrated AI developer has exclusive ownership of B . It conducts research and sells effective AI production services. The developer sells X and purchases the $U = X/B$ units of inference compute needed to supply those services. Because X does not enter the RSI technology in equation (17) directly, its within-date choice can be solved conditional on B . Define optimized operating profit before research expenditure as

$$\pi(B) = \max_{X \geq 0} \left\{ p_X(X)X - \frac{X}{B} \right\}. \quad (18)$$

The notation $\pi(B)$ leaves its dependence on current K, A , and L implicit. The first-order condition equates marginal revenue to the compute cost of one additional service unit:

$$p_X(1 - e_X) = \frac{1}{B}. \quad (19)$$

Lemma 1 in [Appendix A](#) shows that problem (18) has a unique, strictly positive global maximizer characterized by condition (19). Thus, the developer charges the unit cost $1/B$ multiplied by the monopoly markup $1/(1 - e_X)$. The envelope theorem gives

$$\pi_B = \frac{X}{B^2} = \frac{U}{B}. \quad (20)$$

At a fixed X , a marginal increase in AI efficiency saves U/B units of the consumption good.

AI efficiency is nonrival in use. Once B has been developed, using it to augment one unit of inference or research compute does not reduce its ability to augment other units. This does not make effective AI services costless: BU and BM require the rival and costly compute inputs U and M . Nonrivalry also does not rule out excludability. An owner can restrict access to model weights, software, or AI services.

To see the resulting appropriability problem, suppose that competitive service providers had free access to B . Free entry would imply no operating rent. A developer that paid for M would improve a common AI-efficiency index B that competing providers could use without compensating it. The developer would bear the research cost but could not appropriate the resulting benefits. Exclusive ownership allows the developer to capture part of the value created by a higher B and makes the return to research appropriable, as in the standard endogenous-growth treatment of nonrival ideas ([Romer, 1990](#)).

That being said, monopoly is an institutional assumption, not a technological implication of nonrivalry. Licensing, oligopoly, public funding, or open-source provision with research subsidies could also support investment in B (Korinek and Vipra, 2025; Gans, 2024; Demirer et al., 2025). I also abstract from the conflict between general-purpose AI suppliers and application developers studied by Liu and Wan (2026).

3.4.4 The dynamic problem

After the within-date service choice, the developer's net profit is $\Pi_t = \pi(B_t) - M_t$: operating profit less research expenditure. Taking the paths of r, K, A , and L as given and substituting the static solution to problem (18), the developer chooses the path of research compute M to increase AI efficiency B :

$$\max_{\{M_t\}_{t \geq 0}} \int_0^\infty \exp \left\{ - \int_0^t r_s ds \right\} [\pi(B_t) - M_t] dt, \quad (21)$$

subject to the RSI technology given by (17) and taking B_0 as given.

Let q be the current-value shadow price of AI efficiency, measured in units of the current final good per unit of B . An interior choice of research compute M satisfies

$$q\chi\eta B(BM)^{\eta-1}\psi(B) = 1, \quad (22)$$

Equation (22) equates the marginal benefit of research compute to its unit cost. Holding current B fixed, an additional unit of M raises the rate of improvement in AI efficiency by $\chi\eta B(BM)^{\eta-1}\psi(B)$. Multiplication by q values the resulting efficiency gain, accounting for future profit gains and further improvements in AI efficiency.

The current-value costate equation is

$$\frac{\dot{q}}{q} = r - \frac{X}{qB^2} - \eta \frac{\dot{B}}{B} - \dot{B} \frac{\psi'(B)}{\psi(B)}. \quad (23)$$

Rearranging gives the return equation

$$rq = \frac{U}{B} + q\eta \frac{\dot{B}}{B} + q\dot{B} \frac{\psi'(B)}{\psi(B)} + \dot{q}. \quad (24)$$

The first term on the right-hand side of Equation (24) is the inference-compute saving from greater AI efficiency. The second is the research-productivity gain from RSI. With an upper bound, the third term is negative because a higher B leaves less room for further improvement. The final term is the capital gain on AI efficiency. Together these returns must equal the market return rq .

The developer's transversality condition is

$$\lim_{t \rightarrow \infty} \exp \left\{ - \int_0^t r_s ds \right\} q_t B_t = 0. \quad (25)$$

It rules out a path that leaves positive discounted value in the AI-efficiency state.

The preceding conditions are necessary for an interior optimum but do not establish global optimality. Lemma 2 in Appendix A gives sufficient conditions for a trajectory of research expenditure to maximize the present value of the developer's net profit in problem (21).

3.5 Equilibrium

Capital-market clearing identifies household capital with the capital rented by final-good firms. Labor-market clearing gives $L = N$. AI-service market clearing identifies the developer's sales with the final sector's purchases. Combining the household budget, final-good zero profit, developer profit, and market clearing yields the aggregate resource constraint

$$C + \dot{K} + \delta K + U + M = Y. \quad (26)$$

Definition 1 (Equilibrium). *For exogenous initial values $A_0, N_0, K_0, B_0 > 0$, an equilibrium is an infinite-horizon trajectory*

$$\{K_t, B_t, C_t, q_t, L_t, U_t, M_t, X_t, Z_t, Y_t, r_t, w_t, p_{X,t}, s_{X,t}, e_{X,t}, \Pi_t\}_{t \geq 0}$$

such that the developer solves problem (21) and the following conditions hold:

1. *The intratemporal allocation and prices satisfy*

$$L = N, \quad (27a)$$

$$Z = \left[\omega_L (AL)^{(\sigma-1)/\sigma} + \omega_X X^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}, \quad (27b)$$

$$Y = K^\alpha Z^{1-\alpha}, \quad (27c)$$

$$s_X = \frac{\omega_X X^{(\sigma-1)/\sigma}}{\omega_L (AL)^{(\sigma-1)/\sigma} + \omega_X X^{(\sigma-1)/\sigma}}, \quad (27d)$$

$$e_X = \frac{1 - s_X}{\sigma} + \alpha s_X, \quad (27e)$$

$$r = \alpha \frac{Y}{K} - \delta, \quad (27f)$$

$$w = (1 - \alpha)(1 - s_X) \frac{Y}{L}, \quad (27g)$$

$$p_X = (1 - \alpha) s_X \frac{Y}{X}, \quad (27h)$$

$$\frac{1}{B} = p_X (1 - e_X), \quad (27i)$$

$$X = BU, \quad (27j)$$

$$1 = q_X \eta B (BM)^{\eta-1} \psi(B), \quad (27k)$$

$$\Pi = p_X X - U - M. \quad (27l)$$

2. The four dynamic equations are

$$\dot{K} = Y - C - U - M - \delta K, \quad (28a)$$

$$\dot{B} = \chi(BM)^\eta \psi(B), \quad (28b)$$

$$\frac{\dot{C}}{C} = n + r - \rho, \quad (28c)$$

$$\frac{\dot{q}}{q} = r - \frac{X}{qB^2} - \eta \frac{\dot{B}}{B} - \dot{B} \frac{\psi'(B)}{\psi(B)}. \quad (28d)$$

3. The two transversality conditions are

$$\lim_{t \rightarrow \infty} e^{-\rho t} \frac{N_t K_t}{C_t} = 0, \quad (29a)$$

$$\lim_{t \rightarrow \infty} \exp \left\{ - \int_0^t r_s ds \right\} q_t B_t = 0. \quad (29b)$$

Remark 1 (The economy without AI). With $\omega_X = 0$ and $\omega_L = 1$, the model reduces to the Ramsey–Cass–Koopmans economy with logarithmic utility and exogenous labor-augmenting technological progress.

The system contains no separate household budget or final-good zero-profit equation because these are redundant after market clearing. The consolidated income decomposition follows from the equations in the definition:

$$1 = \frac{(r + \delta)K}{Y} + \frac{wL}{Y} + \frac{\Pi}{Y} + \frac{U}{Y} + \frac{M}{Y}. \quad (30)$$

It records how final output is divided among gross capital income, labor income, developer profit, inference compute, and research compute.

4 Equilibrium with an upper bound on AI efficiency

In this section, I characterize the long-run equilibrium behavior for various parameter configurations. The key finding is that long-run output-per-worker growth can exceed the rate of labor-augmenting technological change when the elasticity of substitution between AI production services and labor exceeds one and the upper bound on AI efficiency is sufficiently high.

To identify this threshold, I calculate the net return on capital implied by the production and pricing conditions as AI efficiency approaches its upper bound and labor's income share tends to zero. I compare this return with $\rho + \gamma$, the interest rate that sustains consumption-per-person growth at the exogenous rate γ . When the former exceeds the latter and AI and labor are substitutes, the model supports equilibrium paths along which capital and AI production services grow faster than effective labor, even though AI efficiency converges to a finite limit.

4.1 The return to capital and the efficiency threshold

An upper bound restricts AI efficiency, but not the compute used to supply AI production services. To identify the return when labor's income share vanishes, fix $B > 0$ and consider $s_X \rightarrow 1$. Since $wL/Y = (1 - \alpha)(1 - s_X)$, this is the zero-labor-share limit.

Combine the final-good firm's demand for AI services in equation (13) with the developer's pricing condition in equation (19). The result is:

$$\frac{X}{Y} = B(1 - \alpha)s_X(1 - e_X) \longrightarrow B(1 - \alpha)^2. \quad (31)$$

The limit uses $e_X \rightarrow \alpha$ from Equation (15). The CES aggregator in equation (8) and the definition of s_X in equation (10) imply

$$\frac{Z}{X} = \left(\frac{\omega_X}{s_X} \right)^{\sigma/(\sigma-1)} \longrightarrow \omega_X^{\sigma/(\sigma-1)}. \quad (32)$$

Using $Z/Y = (Z/X)(X/Y)$ and the limits in equations (31) and (32), I obtain

$$\frac{Y}{K} = \left(\frac{Z}{Y} \right)^{(1-\alpha)/\alpha} \longrightarrow \left[(1 - \alpha)^2 B \omega_X^{\sigma/(\sigma-1)} \right]^{(1-\alpha)/\alpha}. \quad (33)$$

Let $\mathcal{R}(B)$ denote the net return on capital in this zero-labor-share limit at a given B . Substituting the limiting output-capital ratio in equation (33) into $r = \alpha Y/K - \delta$ gives

$$\mathcal{R}(B) \equiv \alpha \left[(1 - \alpha)^2 B \omega_X^{\sigma/(\sigma-1)} \right]^{(1-\alpha)/\alpha} - \delta. \quad (34)$$

This is a return at a particular limit of the production and pricing conditions, not necessarily an equilibrium interest rate. For $\sigma > 1$, $s_X \rightarrow 1$ means that AI production services become arbitrarily large relative to effective labor. For $\sigma < 1$, $s_X \rightarrow 1$ means that effective labor services become arbitrarily large relative to AI production services. Whether the equilibrium interest rate converges to $\mathcal{R}(\bar{B})$ depends on both σ and $\bar{B}$, as shown below.

By the Euler equation (5), consumption per person grows at $r - \rho$. If this growth rate converges to γ , the interest rate must converge to $\rho + \gamma$. Define $\mathcal{B}(\sigma)$ as the AI-efficiency level at which the zero-labor-share return equals this benchmark:

$$\begin{aligned} \mathcal{R}(\mathcal{B}(\sigma)) &= \rho + \gamma, \\ \mathcal{B}(\sigma) &= \frac{[(\rho + \gamma + \delta)/\alpha]^{\alpha/(1-\alpha)}}{(1 - \alpha)^2 \omega_X^{\sigma/(\sigma-1)}}. \end{aligned} \quad (35)$$

The inequalities between $\bar{B}$ and $\mathcal{B}(\sigma)$ –whether AI can attain sufficient efficiency– and between σ and 1 –whether labor and AI are substitutes or complements– distinguish the long-run equilibrium regimes characterized below. As [Duffy and Papageorgiou \(2000\)](#) explain for CES models of capital

and labor, an elasticity above one permits sustained growth only if the limiting return on capital is sufficiently high.

4.2 Equilibrium existence and long-run regimes

Define $g_{Y/N}$ as the growth rate of output per worker and g_w as the wage growth rate. The next proposition characterizes the long-run behavior of equilibrium trajectories.

Proposition 1 (Equilibrium existence and long-run characterization). *Fix $\bar{B} < \infty$. For each parameter configuration covered by a case below, there exists a nonempty open set of initial conditions (A_0, N_0, K_0, B_0) such that every initial state in that set has an equilibrium trajectory with $B_t \rightarrow \bar{B}$ as $t \rightarrow \infty$ and the following properties:*

(i) **Complementarity.** *If $\sigma < 1$ and $\bar{B} > \mathcal{B}(\sigma)$, then*

$$g_{Y/N} \rightarrow \gamma, \quad g_w \rightarrow \gamma, \quad r \rightarrow \rho + \gamma, \quad \lim_{t \rightarrow \infty} s_X \in (0, 1). \quad (36)$$

(ii) **Unit elasticity.** *If $\sigma = 1$, then*

$$g_{Y/N} \rightarrow \gamma, \quad g_w \rightarrow \gamma, \quad r \rightarrow \rho + \gamma, \quad s_X = \omega_X \quad (37)$$

(iii) **Substitution with a labor bottleneck.** *If $\sigma > 1$ and $\bar{B} < \mathcal{B}(\sigma)$, then*

$$g_{Y/N} \rightarrow \gamma, \quad g_w \rightarrow \gamma, \quad r \rightarrow \rho + \gamma, \quad \lim_{t \rightarrow \infty} s_X \in (0, 1). \quad (38)$$

(iv) **AI-dominated growth.** *If $\sigma > 1$ and $\bar{B} > \mathcal{B}(\sigma)$, then*

$$g_{Y/N} \rightarrow \mathcal{R}(\bar{B}) - \rho > \gamma, \quad (39)$$

$$g_w \rightarrow \left(1 - \frac{1}{\sigma}\right) \gamma + \frac{1}{\sigma} (\mathcal{R}(\bar{B}) - \rho) > \gamma, \quad (40)$$

$$r \rightarrow \mathcal{R}(\bar{B}), \quad s_X \rightarrow 1.$$

Proof. See Appendix A, [Proof of Proposition 1](#). □

I omit the case $0 < \sigma < 1$ with $\bar{B} < \mathcal{B}(\sigma)$ from Proposition 1 because it describes an AI production bottleneck induced by a sufficiently low upper bound on AI efficiency. The [corresponding result in the digital edition](#) shows that, in any equilibrium in this region, AI services become scarce relative to effective labor and long-run average output-per-worker growth is below the exogenous rate γ of labor-augmenting technological progress.

The labor-bottleneck cases of Proposition 1 show that autonomous AI research need not eliminate the production bottleneck, a constraint also emphasized by [Aghion et al. \(2019\)](#). AI efficiency improves in each of these cases, including complementarity. In the complementary labor-bottleneck

equilibria, expanding AI services cannot substitute for the continued expansion of effective labor. As AI-efficiency growth tends to zero, output per worker and wages therefore return to growth at γ . A positive limiting labor share and the absence of a permanent growth acceleration are two consequences of this production bottleneck. Substitution is needed for a permanent increase in growth under an upper bound, not for research to improve AI efficiency. Moreover, $\sigma > 1$ is not sufficient to escape the labor bottleneck: if $\mathcal{R}(\bar{B}) < \rho + \gamma$, the return at a vanishing labor share is too low to sustain growth faster than effective labor.

The AI-dominated regime has an asymptotically linear production structure, as in the capital-accumulation mechanisms of [Rebelo \(1991\)](#) and [Jones and Manuelli \(1990\)](#). Equation (33) gives a positive limiting output–capital ratio even as B approaches the upper bound. Capital and inference compute can expand together, sustaining output-per-worker growth above γ even when B is arbitrarily close to $\bar{B}$ and further AI-efficiency gains vanish, as in the compute-expansion mechanism of [Restrepo \(2025\)](#). RSI can help the economy reach an efficiency at which this mechanism supports growth faster than γ . However, continuing efficiency improvements are not necessary to sustain it. Section 5 shows that sufficiently high initial efficiency also allows this growth mechanism to operate under competition without RSI.

This joint expansion also explains the higher long-run interest rate. In the labor-bottleneck regime, the interest rate converges to $\rho + \gamma$. In the AI-dominated regime, however, AI services expand alongside capital, so production need not be constrained by the growth of effective labor. The output–capital ratio converges to a level at which $r = \alpha Y/K - \delta$ remains above $\rho + \gamma$: accumulating more capital does not restore the labor-bottleneck return.

The proof of Proposition 1 first removes the long-run trends from growing quantities and shrinking efficiency gaps. For each parameter configuration covered by the proposition, I identify candidate stationary values of these normalized variables. I then construct equilibrium trajectories from every sufficiently nearby initial state and verify convergence to those values. These analytical results guide the algorithm’s boundary conditions, as described in Appendix B.

Convergence to the upper bound is not guaranteed from arbitrary initial states. Under complementarity and positive depreciation, sufficiently low initial AI efficiency can make output insufficient to replace depreciating capital. If the initial capital stock is also sufficiently low, the resources available for research cannot raise AI efficiency enough to reverse the decline in capital. This obstacle can arise even when $\bar{B} > \mathcal{B}(\sigma)$, as shown in the counterexample in the [digital edition](#).

Remark 2 (Wages and AI revenue in the AI-dominated regime). In the AI-dominated equilibria of Proposition 1(iv), output per worker grows asymptotically faster than wages, even though both grow faster than labor-augmenting technological progress:

$$\lim_{t \rightarrow \infty} g_{Y/N} > \lim_{t \rightarrow \infty} g_w > \gamma.$$

Labor’s income share therefore converges to zero, while the AI industry’s revenue as a share of output converges to $1 - \alpha$.

Figure 3 shows where the long-run regimes characterized in Proposition 1 lie relative to $\mathcal{B}(\sigma)$. For $0 < \sigma < 1$, the equilibria above the threshold have output-per-worker growth converging to γ and a positive limiting labor income share. For $\sigma > 1$, these same long-run outcomes occur below the threshold. Above it, the regime is AI-dominated: output per worker grows asymptotically faster than γ and labor's income share converges to zero.

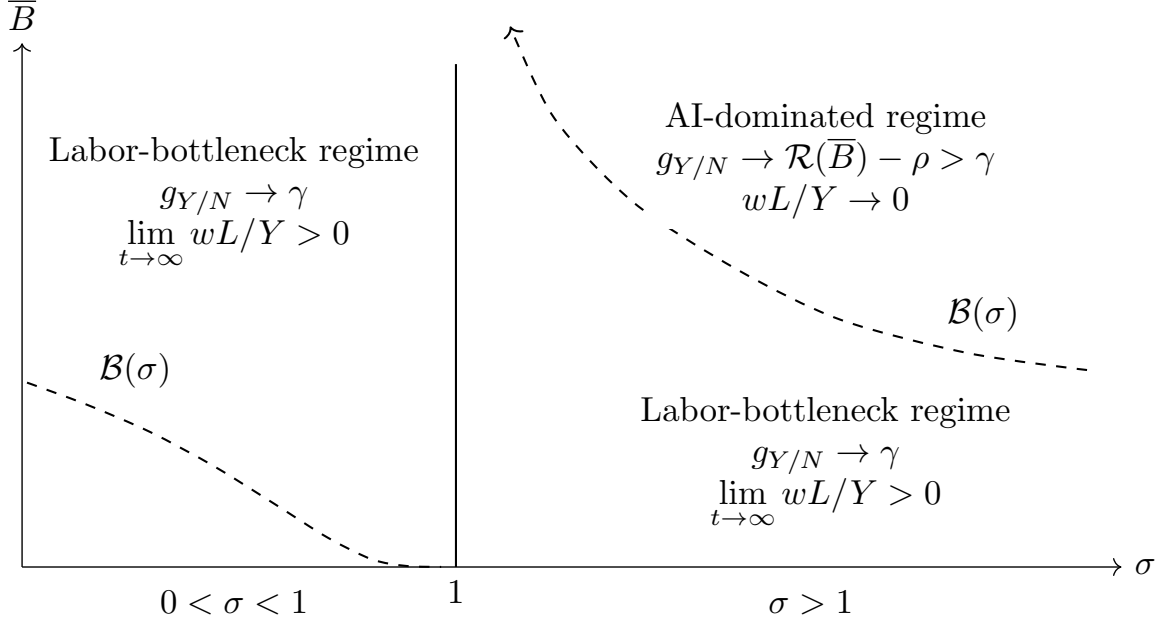


Figure 3: Growth regimes with an upper bound on AI efficiency.

Table 1 summarizes the four cases of Proposition 1. For $\sigma > 1$, an upper bound below $\mathcal{B}(\sigma)$ makes the return on capital at a vanishing labor share too low for capital and AI services to outgrow effective labor. Above the threshold, that return supports their faster growth along the AI-dominated equilibria. Across these equilibrium regimes, long-run wage growth exceeds γ if and only if labor's income share converges to zero.

Table 1: Long-run equilibrium characterization

Elasticity	AI efficiency bound	$\lim g_{Y/N}$	$\lim g_w$	$\lim wL/Y$	$\lim r$
$0 < \sigma < 1$	$\bar{B} > \mathcal{B}(\sigma)$	γ	γ	Positive	$\rho + \gamma$
$\sigma = 1$	Any finite $\bar{B}$	γ	γ	$(1 - \alpha)\omega_L$	$\rho + \gamma$
$\sigma > 1$	$\bar{B} < \mathcal{B}(\sigma)$	γ	γ	Positive	$\rho + \gamma$
$\sigma > 1$	$\bar{B} > \mathcal{B}(\sigma)$	$\mathcal{R}(\bar{B}) - \rho$	$(1 - \frac{1}{\sigma})\gamma + (\frac{1}{\sigma})(\mathcal{R}(\bar{B}) - \rho)$	Zero	$\mathcal{R}(\bar{B})$

4.3 Substitution and wage growth

When AI and labor are substitutes, cheaper AI reduces labor demand at a given wage and budget for labor and AI services. But this fixed-budget comparison misses the expansion of production in equilibrium. With sufficiently efficient AI, capital and AI services can grow faster than effective

labor while maintaining a return high enough to sustain capital accumulation. Because substitution is imperfect, this joint expansion raises workers' marginal products and sustains wage growth above the rate of labor-augmenting technological progress, even as labor's income share vanishes. With complementarity, by contrast, additional AI cannot compensate sufficiently for effective labor growing more slowly. The resulting decline in the return to capital prevents that faster expansion from continuing indefinitely. Wage growth in the labor-bottleneck regime therefore returns to the rate of labor-augmenting technological progress as AI-efficiency gains vanish.

To see why wages accelerate even as labor's share vanishes, I rewrite the wage condition (12) using the CES aggregator in equation (8):

$$w = (1 - \alpha)\omega_L A \left(\frac{K}{Z}\right)^\alpha \left(\frac{Z}{AL}\right)^{1/\sigma}. \quad (41)$$

In the AI-dominated regime, capital and AI services expand together, so K/Z converges to a positive constant while $Z/(AL)$ keeps growing. Although labor accounts for a declining share of output, its marginal product grows faster than A : workers participate in an expanding production process and remain imperfectly replaceable by AI.

Equations (39) and (40) make the distribution of this growth acceleration explicit:

$$\lim_{t \rightarrow \infty} g_w = \gamma + \frac{1}{\sigma} \left(\lim_{t \rightarrow \infty} g_{Y/N} - \gamma \right). \quad (42)$$

For $1 < \sigma < \infty$, only a fraction $1/\sigma$ of output-per-worker growth above γ translates into wage growth above γ . For example, long-run output-per-worker growth of 4.0%, with $\gamma = 1.0\%$ and $\sigma = 1.5$, implies wage growth of 3.0%. This wage-growth mechanism has a direct antecedent in the CES example of Palivos and Karagiannis (2010).

A higher elasticity has two opposing effects on long-run wage growth. On one hand, it raises the limiting return in equation (34) and therefore output-per-worker growth in the AI-dominated regime. On the other, it reduces the fraction $1/\sigma$ of the additional growth that reaches wages. This hints at the existence of an elasticity of substitution that maximizes long-run wage growth.

Proposition 2 (Substitution and wage growth). *If $\bar{B} > \mathcal{B}(\sigma)$ for some $\sigma > 1$, then long-run wage growth is maximized at a unique finite elasticity of substitution $\sigma^* > 1$.*

Proof. See Appendix A, Proof of Proposition 2. □

The association between permanently faster wage growth and a vanishing labor share depends on bounded AI efficiency. In Appendix C, I show that, without this bound, continued AI improvement at unit elasticity can sustain a common growth rate for wages and output per worker above γ , while preserving a constant positive labor share.

5 Competitive Equilibrium

I now replace the monopolistic developer with competitive AI suppliers. This comparison separates the incentive to improve AI from the use of technology already developed. I focus on $\sigma > 1$, where both labor-bottleneck and AI-dominated growth are possible. I first characterize the competitive equilibrium and then compare it with the monopoly equilibria in Section 4.

5.1 Characterization

Consider now the case where multiple AI suppliers have free access to the same efficiency B . Entry is free, and any improvement becomes immediately available to every supplier, without royalties or other compensation to its developer. There are no research subsidies or publicly financed research. Suppliers, households and firms take prices as given.

Competition sets $p_X = 1/B$. Since $X = BU$, revenue then exactly covers inference compute: $p_X X = U$. A supplier that invests in research bears its cost but earns neither operating profit nor royalties on the improvement. It therefore chooses $M = 0$, leaving $B = B_0$. Final-good firms continue to solve problem (9), with production given by equations (7)–(8). Their demand for AI services in equation (13) then implies $U = (1 - \alpha)s_X Y$, since $X = B_0 U$. The household continues to choose consumption and saving according to the Euler equation (5) and the transversality condition (6). Capital accumulation follows the resource constraint (26) with $M = 0$. Competitive suppliers' optimality replaces the monopolist's pricing, research, and shadow-price conditions.

As labor's income share vanishes, the net return on capital under marginal-cost pricing approaches

$$\mathcal{R}^{\text{mc}}(B) = \alpha \left[(1 - \alpha) B \omega_X^{\sigma/(\sigma-1)} \right]^{(1-\alpha)/\alpha} - \delta. \quad (43)$$

Let $\mathcal{B}^{\text{mc}}(\sigma)$ be the unique efficiency at which $\mathcal{R}^{\text{mc}}(\mathcal{B}^{\text{mc}}(\sigma)) = \rho + \gamma$. Unlike the monopoly economy, the competitive economy cannot improve from B_0 toward $\bar{B}$. Its long-run regime depends on the efficiency inherited at the initial date.

Proposition 3 (Competitive equilibrium). *Suppose $\sigma > 1$. A unique competitive equilibrium exists from every admissible initial state (A_0, N_0, K_0, B_0) . At every date,*

$$M = 0, \quad B = B_0, \quad p_X = 1/B_0, \quad \Pi = 0. \quad (44)$$

Its long-run outcomes on either side of the threshold are:

(i) *If $B_0 < \mathcal{B}^{\text{mc}}(\sigma)$, labor remains a bottleneck:*

$$g_{Y/N} \rightarrow \gamma, \quad g_w \rightarrow \gamma, \quad r \rightarrow \rho + \gamma, \quad \lim_{t \rightarrow \infty} wL/Y > 0.$$

(ii) If $B_0 > \mathcal{B}^{\text{mc}}(\sigma)$, growth is AI-dominated:

$$\begin{aligned} g_{Y/N} &\rightarrow \mathcal{R}^{\text{mc}}(B_0) - \rho > \gamma, \\ g_w &\rightarrow \left(1 - \frac{1}{\sigma}\right) \gamma + \frac{1}{\sigma} (\mathcal{R}^{\text{mc}}(B_0) - \rho) > \gamma, \\ r &\rightarrow \mathcal{R}^{\text{mc}}(B_0), \quad wL/Y \rightarrow 0. \end{aligned} \tag{45}$$

Proof. See Appendix A, Proof of Proposition 3. $\square$

Neither monopoly nor RSI is necessary to sustain AI-dominated growth. At fixed efficiency, $X = B_0 U$ can continue expanding through greater compute use. When B_0 exceeds the competitive threshold $\mathcal{B}^{\text{mc}}(\sigma)$, capital and AI services grow together faster than effective labor. This is the capital-accumulation mechanism identified in Section 4.2, now sustained without private research. Below the threshold, diminishing returns bring the interest rate back to $\rho + \gamma$.

5.2 Comparison with monopoly

I now compare the efficiency required under the two pricing rules, isolating the effect of pricing from the difference in research incentives.

Proposition 4 (Monopoly pricing and the AI-efficiency threshold). *Suppose AI and labor are substitutes ($\sigma > 1$). Let $\mathcal{B}(\sigma)$ and $\mathcal{B}^{\text{mc}}(\sigma)$ denote the AI-efficiency thresholds under monopoly pricing and marginal-cost pricing, respectively. These thresholds identify the efficiency needed for the return on capital to reach $\rho + \gamma$ in the limit where labor's income share vanishes. Monopoly pricing requires a higher AI efficiency:*

$$\mathcal{B}^{\text{mc}}(\sigma) = (1 - \alpha)\mathcal{B}(\sigma) < \mathcal{B}(\sigma). \tag{46}$$

Proof. See Appendix A, Proof of Proposition 4. $\square$

To understand why monopoly pricing raises the threshold, compare the same AI technology with the same capital and effective labor under the two pricing rules. At marginal cost, AI services are cheaper, so final-good firms use more of them and produce more with the same capital. This raises the output-capital ratio and therefore the return on capital. For the threshold comparison, evaluate both returns as labor's income share tends to zero. In this limit, monopoly pricing requires a higher AI efficiency for the return on capital to reach the benchmark $\rho + \gamma$. With $\alpha = 0.33$, marginal-cost pricing reaches this benchmark at only 67.0% of the AI efficiency required under monopoly pricing.

Figure 4 adds the competitive threshold to the substitution region of Figure 3. Efficiencies between the curves exceed the competitive threshold but remain below the monopoly threshold.

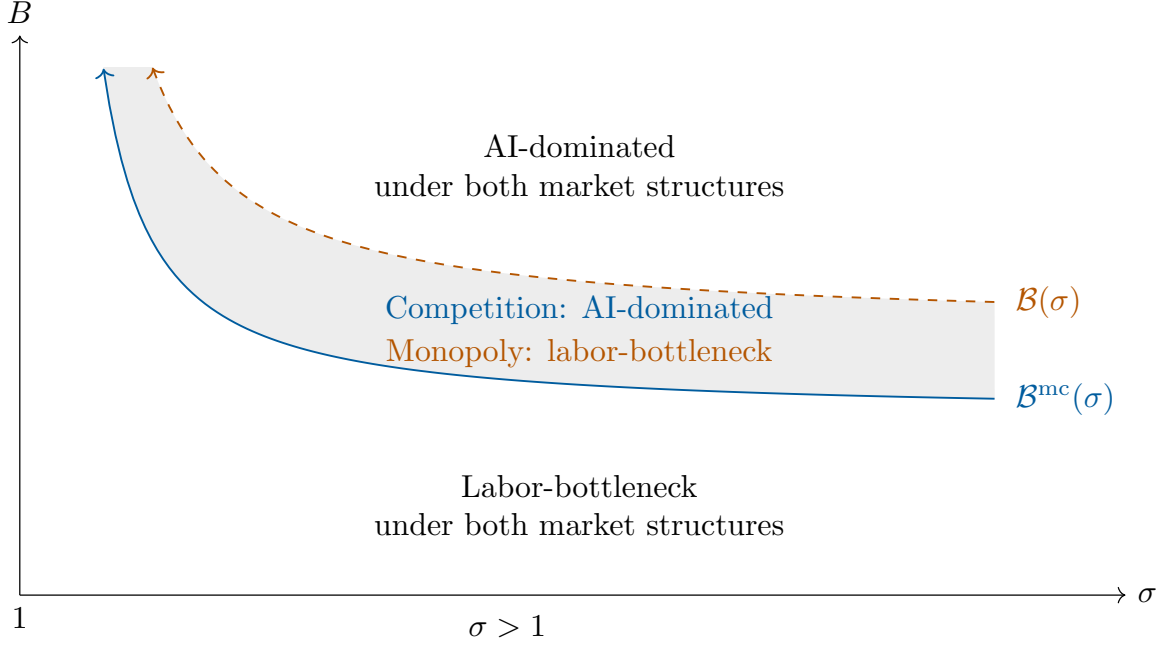


Figure 4: Long-run regimes under monopoly and competition ($\sigma > 1$). The blue solid curve is the competitive threshold; the orange dashed curve is the monopoly threshold. Their relationship is $\mathcal{B}^{\text{mc}}(\sigma) = (1 - \alpha)\mathcal{B}(\sigma)$ (Proposition 4). Read the vertical axis as B_0 under competition and as $\bar{B}$ under monopoly.

Corollary 1 (Monopoly and the labor bottleneck). *Suppose $\sigma > 1$ and*

$$\mathcal{B}^{\text{mc}}(\sigma) < \bar{B} < \mathcal{B}(\sigma). \quad (47)$$

There is a set of initial states with $\mathcal{B}^{\text{mc}}(\sigma) < B_0 < \bar{B}$ from which the competitive equilibrium converges to AI-dominated growth, while a monopoly equilibrium converges to the labor-bottleneck regime. AI efficiency remains at B_0 under competition but approaches the higher level $\bar{B}$ under monopoly.

Proof. See Appendix A, [Proof of Corollary 1](#). □

In this case, the efficiency gains generated by monopoly-led research do not compensate for the reduction in AI use caused by monopoly pricing. Competition sustains permanently faster output-per-worker and wage growth despite retaining a less efficient AI technology. AI services per unit of effective labor converge to a positive constant under monopoly but grow without bound under competition.

Corollary 2 (Growth under competition and monopoly). *Suppose $\sigma > 1$. Consider competitive and monopoly equilibria with the same parameters and initial stocks, both converging to AI-dominated growth, with AI efficiency approaching $\bar{B}$ under monopoly. Monopoly delivers a higher long-run return to capital and faster long-run growth in output per worker and wages if and only if*

$$B_0 < (1 - \alpha)\bar{B}. \quad (48)$$

Competition delivers higher limiting returns and growth rates when the inequality is reversed.

Proof. See Appendix A, Proof of Corollary 2. □

There is therefore no universal growth ranking between competition and monopoly. Conditional on the existence of the compared equilibria, monopoly-led research can allow an escape from the labor bottleneck when $B_0 < \mathcal{B}^{\text{mc}}(\sigma)$ but $\bar{B} > \mathcal{B}(\sigma)$. Corollary 1 establishes the opposite regime ranking, while Corollary 2 shows what determines the ranking when both economies are AI-dominated. The relevant comparison is not efficiency alone, but whether monopoly's efficiency gains compensate for the reduction in AI use caused by its markup.

5.3 Research incentives and policy

At given AI efficiency, capital, and effective labor, marginal-cost pricing maximizes $Y - X/B$ and leaves more resources available for consumption or capital accumulation than monopoly pricing. However, the competitive arrangement studied here also removes the private return to research. This raises a classic time-inconsistency problem (Kyddland and Prescott, 1977). A government may promise monopoly protection to encourage research, but later wish to remove the markup once the technology has been developed. Anticipating this policy reversal can weaken the developer's incentive to invest in research.

An announced rule that ends monopoly protection at a specified efficiency level is different from renegeing on a promise. It nevertheless changes the developer's problem. Without compensation, reaching that level would remove future monopoly rents, potentially giving the developer an incentive to delay reaching it. The threshold $\mathcal{B}^{\text{mc}}(\sigma)$ identifies the AI efficiency at which the limiting return on capital under marginal-cost pricing reaches $\rho + \gamma$. It does not identify the optimal time to end monopoly protection. Choosing that time requires accounting for the research response and comparing consumption paths under alternative rules.

Patent buyouts offer a possible way to reward innovation without retaining the monopoly markup. Kremer (1998) studies government purchases of patents that are generally placed in the public domain. Applied here, the idea would be to compensate the developer for the rights to an existing AI technology and allow competing suppliers to use it. Competitive supply at the same efficiency would require access to the algorithms and model weights. The buyout would also need a credible valuation and payment rule, a source of financing, and arrangements to reward subsequent improvements. Incentives for further development depend on whether developers can appropriate

its benefits. These issues are particularly important for RSI, where today’s efficiency is also an input into tomorrow’s research. In this paper I leave aside the design of such a policy, including its effects on research and welfare.

6 Numerical simulations

The exercises in this section answer two questions. First, how does an economy adjust when RSI becomes available under an existing monopoly, and how does that adjustment depend on research productivity? Second, can research financed by monopoly rents compensate for the restriction on AI use caused by monopoly pricing and enable permanently faster growth? The [supplementary simulations in the digital edition](#) also examine the opposite outcome: can monopoly pricing eliminate AI-dominated growth despite the efficiency gains from research? Appendix B describes the algorithm used to find the equilibrium and presents the links to replication files.

6.1 Parameters and initial conditions

In the next two subsections, I compare the transitional dynamics across four parameter configurations corresponding to the cases in Proposition 1. I study an unanticipated activation of RSI at date zero, retaining monopoly pricing before and after the event. Before date zero, the economy already uses AI production services, but AI efficiency is fixed at B_0 . For the chosen value of B_0 , each economy starts on a labor-bottleneck balanced growth path. I compute this path with the research productivity parameter χ equal to zero to obtain the corresponding initial capital stock K_0 for each parameter configuration. Setting $\chi = 0$ rules out improvements in AI efficiency, not the expansion of AI services through inference compute U . By treating activation as unanticipated, I exclude the advance adjustment of consumption and saving emphasized in the technology-news literature reviewed by [Beaudry and Portier \(2014\)](#).

I compare four economies with $\sigma = 0.9, 1.0, 1.1, 1.5$, holding all other parameters and the initial values A_0 , N_0 , and B_0 fixed. For each economy, I choose K_0 to place it on its pre-RSI balanced growth path. In particular, for each elasticity, I choose the pre-event capital stock to satisfy

$$r_0 = \rho + \gamma, \quad \frac{K_0}{Y_0} = \frac{\alpha}{\rho + \gamma + \delta} = 3.30. \quad (49)$$

This setup guarantees that Y , X , U , w , r , and p_X behave continuously rather than jumping at date zero. Consumption, research expenditure, and growth rates can nevertheless jump when RSI becomes available. Within each elasticity, the two research-productivity scenarios inherit exactly the same stocks; consumption and research are chosen anew after activation.

Table 2 reports the parameter values and initial stocks used for the simulations. Time is measured in years, and the four values of K_0 follow the order of the elasticities in the table. I use standard values for α , δ , and ρ , while n and γ are guided by long-run population and productivity projections. The AI production weight $\omega_X = 0.10$ and research exponent $\eta = 0.20$ are illustrative

choices, not empirical estimates. I choose them to satisfy $\omega_X + \eta < 1$ and $\eta < \alpha$.

The common upper bound on AI efficiency satisfies $\mathcal{B}(1.5) < \bar{B} < \mathcal{B}(1.1)$. Thus, AI and labor are substitutes in both the $\sigma = 1.1$ and $\sigma = 1.5$ cases, but only the latter converges to the AI-dominated regime. Holding this upper bound fixed, I study two levels of research productivity for each of the four elasticities of substitution: $\chi = 7.5$ in the higher-productivity scenario and $\chi = 1.5$ in the lower-productivity scenario. For a given elasticity, both lead to the same long-run regime, output and wage growth rates, and interest rate, but produce different initial adjustments and transition speeds. These choices are intended to show how research productivity affects the short-run impact of RSI and the timing of subsequent growth.

Table 2: Simulation parameters and initial stocks

Symbol	Value	Interpretation	Source
α	0.33	Capital share	Standard calibration value.
δ	0.05	Depreciation rate	Standard calibration value.
ρ	0.04	Discount rate	Standard calibration value.
n	0.003	Population growth rate	Rounded constant-rate approximation to the UN 2024–2100 world projection (United Nations, 2024).
γ	0.01	Growth of labor-augmenting technology	OECD labor-productivity projections (OECD, 2025).
η	0.20	Research exponent	Illustrative.
ω_X	0.10	AI production weight	Illustrative.
ω_L	0.90	Effective-labor weight	$1 - \omega_X$.
χ	7.5 (higher); 1.5 (lower)	Research productivity after activation	Illustrative.
σ	0.90; 1.00; 1.10; 1.50.	Elasticity of substitution	Illustrative.
$\bar{B}$	1,361	AI-efficiency upper bound	$1.10 \times \mathcal{B}(1.50)$.
A_0	1	Initial labor-efficiency units	Normalization.
N_0	1	Initial population	Normalization.
B_0	13.6	Initial AI efficiency	$0.01 \times \bar{B}$.
K_0	3.43; 4.65; 5.04; 5.50.	Initial capital	Adjusted to match $K_0/Y_0 = 3.30$, the pre-RSI balanced growth path ratio.

6.2 Fast takeoff

I first ask how research, consumption, and capital accumulation respond when RSI becomes available under an existing monopoly, and how this adjustment differs across economies approaching the two long-run regimes. I compare the four elasticities under relatively high research productivity, $\chi = 7.5$. Research productivity affects the transition, but not the limiting growth rates characterized in

Section 4. All four economies inherit the pre-RSI balanced growth paths described in Section 6.1. Table 2 reports their common parameters and elasticity-specific initial capital stocks. I examine quantities first, then wages and prices, and finally income shares.

I use Figure 5 to present the annual growth rates of output, AI production services, and capital per unit of effective labor: $Y/(AL)$, $X/(AL)$, and $K/(AL)$. Zero therefore means that the corresponding aggregate grows at the same rate as effective labor, $n + \gamma$. The upper row shows two pre-event years and the first ten years after RSI activation. The lower row follows the same variables from year 10 to year 500. Throughout the three figures, green dashed, black solid, orange dotted, and blue dash-dotted lines denote $\sigma = 0.90, 1, 1.10, 1.50$, respectively. The vertical line marks RSI activation, and black dotted horizontal lines mark the analytical limits for $\sigma = 1.50$.

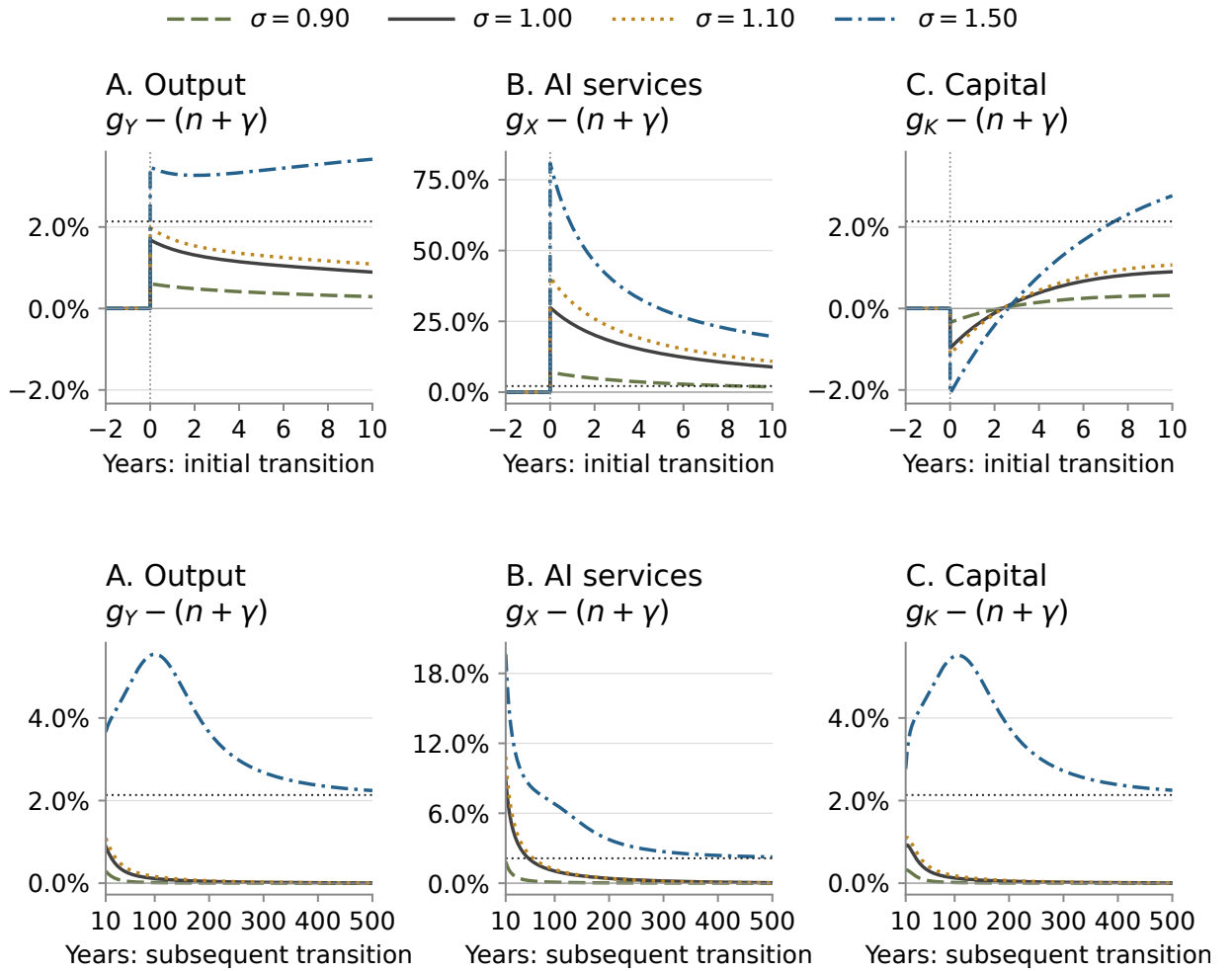


Figure 5: Higher research productivity ($\chi = 7.5$): growth of output, AI production services, and capital per unit of effective labor. The upper row shows two pre-event years and years 0–10; the lower row shows years 10–500. The vertical line marks RSI activation.

The top row of Figure 5 shows the short-run effects of RSI activation. At $\sigma = 1.50$, $X/(AL)$

initially grows at 80.8% per year, whereas $Y/(AL)$ grows at 3.5%. Research begins to raise B , allowing more services to be obtained from compute through $X = BU$. Initial growth of $Y/(AL)$ is smaller at the other elasticities, reaching only 0.6% at $\sigma = 0.90$. The third panel shows the capital response: capital per unit of effective labor, $K/(AL)$, falls in every case. Households increase consumption at RSI activation while the developer starts spending on research. The resource constraint thus requires lower investment in physical capital. Capital accumulation subsequently recovers as output expands and as the marginal product of capital increases.

The bottom row of Figure 5 separates temporary growth gains from permanent ones. At $\sigma = 0.90, 1, 1.10$, all three normalized growth rates decline toward zero. This is the labor-bottleneck outcome in Proposition 1: output per worker eventually grows at $\gamma = 1.0\%$. At $\sigma = 1.50$, output and capital growth instead rise further before declining toward their long-run values. As explained in Section 4.2, sufficient substitution and a sufficiently high efficiency bound allow capital and inference compute to expand together without being limited by effective-labor growth.

I next examine wages and prices in Figure 6. The figure shows annual wage growth g_w , the annual net interest rate r , and the level of p_X in final-good units on a logarithmic scale. The top panels present the short-run dynamics and the bottom panels present the long-run dynamics.

Wage growth increases in all four economies, but the initial response is not ordered monotonically by substitutability: the immediate increase is larger at $\sigma = 1$ than at $\sigma = 1.50$. Wage growth in the highest-elasticity case soon surpasses the other cases. After the initial jump, wage growth declines in the labor-bottleneck cases but continues to rise in the AI-dominated case.

The interest rate starts at 5.0% in every case and then rises as output initially outpaces capital. The higher return accompanies faster consumption-per-person growth, as required by the Euler condition in Equation (5). A related mechanism underlies Chow et al. (2026), who link faster expected consumption growth to higher long-term real interest rates.

The bottom row of Figure 6 shows the long-run adjustment of wage growth, the interest rate, and the AI-service price. The three labor-bottleneck economies return toward wage growth of 1.0% and an interest rate of 5.0%. At $\sigma = 1.50$, both the wage growth and the interest rate rise above their long-run values before declining. Equation (40) gives limiting wage growth of 2.4%, below output-per-worker growth of 3.1% but above labor-augmenting technological progress of 1.0%. Meanwhile, the interest rate tends to $\mathcal{R}(\bar{B}) = 7.1\%$, above its pre-RSI level.

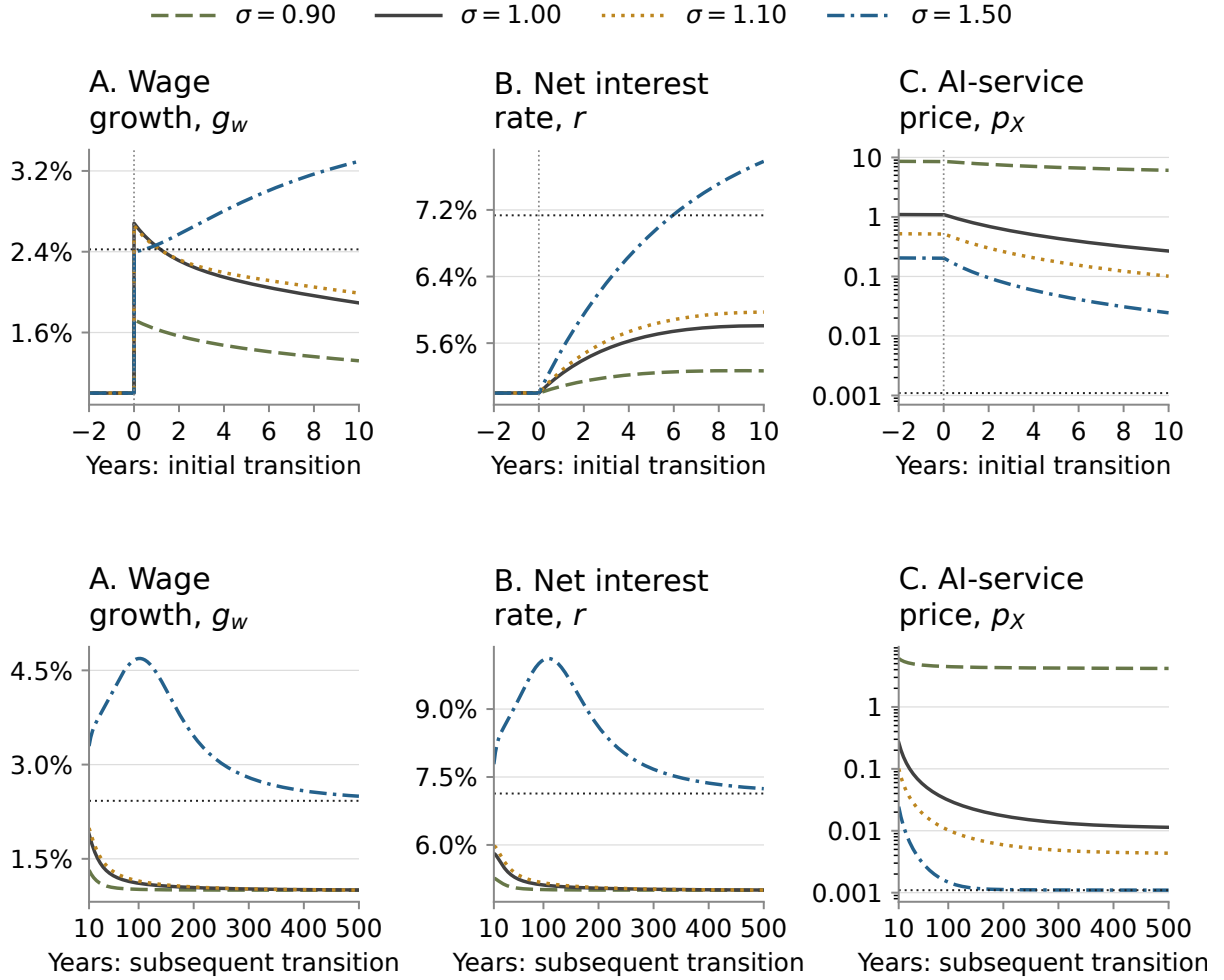


Figure 6: Higher research productivity ($\chi = 7.5$): wage growth, the net interest rate, and the AI-service price. The price is in final-good units on a logarithmic scale. The vertical line marks RSI activation.

Finally, I use Figure 7 to show labor income wL/Y , developer net profit Π/Y , inference expenditure U/Y , and research expenditure M/Y , all as a share of output. The top four panels show the short run and the bottom four panels show the long run. Line styles again identify the four elasticities. The omitted gross capital income share remains $\alpha = 33.0\%$.

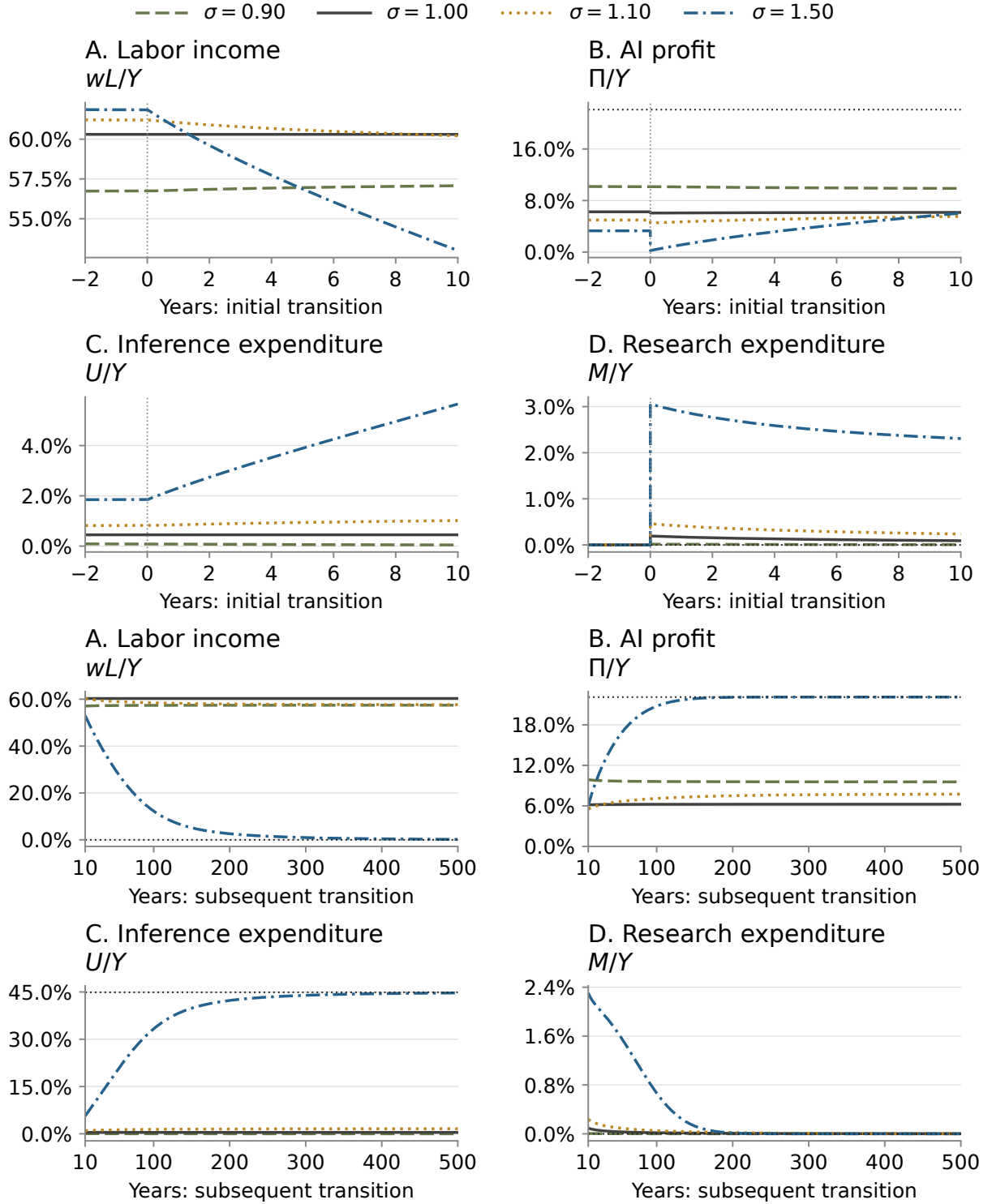


Figure 7: Higher research productivity ($\chi = 7.5$): labor income, AI net profit, inference expenditure, and research expenditure relative to output. The first two rows show the pre-event BGP and years 0–10; the last two show years 10–500. The omitted gross capital share remains $\alpha = 0.33$.

In the short run, research spending as a share of output jumps in all four economies. The

adjustment is strongest at $\sigma = 1.50$: immediately after activation, research absorbs 59.4% of AI revenue, compared with 35.9% for inference compute. The developer’s net profit therefore falls sharply. Profit subsequently recovers as research absorbs a smaller share of revenue and AI sales expand. Labor’s income share also falls most strongly in this scenario, while changes are much smaller in the other three economies.

Over the longer horizon, the AI-dominated economy combines permanently faster wage growth with an even faster expansion of output per worker, so labor’s income share tends to zero. Rising wages alongside a declining labor share also appear in [Autor and Kausik \(2026\)](#) and [Ray and Mookherjee \(2022\)](#). As labor’s share falls, AI revenue rises relative to output. Inference compute eventually absorbs 67.0% of AI revenue, while research expenditure vanishes relative to that revenue.

6.3 Slow takeoff

I next ask whether lower research productivity makes the early transition less informative about the eventual growth regime, and how it changes the initial allocation of resources. I set $\chi = 1.5$, one fifth of its value in Section 6.2. All other parameters and initial stocks remain unchanged, including the elasticity-specific pre-RSI capital stocks in Table 2. Consumption and research are solved again for each economy. The new exercise changes both the initial adjustment and the timing of the transition, but not the limiting growth rates, interest rates, and income shares.

I use Figure 8 to present the growth of $Y/(AL)$, $X/(AL)$, and $K/(AL)$. As before, zero means growth at the rate of effective labor, $n + \gamma$. The upper row shows two pre-event years and the first ten years after RSI activation. The lower row follows the same trajectories from year 10 to year 500. Throughout the three figures, line styles identify the same four elasticities as in Figure 5, and black dotted lines mark the analytical limits for $\sigma = 1.50$.

The upper row shows a much smaller initial increase in growth. Growth of $Y/(AL)$ ranges from 0.1% to 0.5% across the four economies. AI services still grow faster than output, but their initial normalized growth rate at $\sigma = 1.50$ is 11.6%, compared with 80.8% in the preceding exercise. Lower research productivity and a smaller research allocation produce slower efficiency improvements. Consumption still rises at activation, and research uses resources that could otherwise finance capital accumulation. Both adjustments are smaller, however, so $K/(AL)$ falls less. As capital accumulation recovers, output growth rises during the first decade even though AI-service growth is already slowing. The four economies therefore begin with more moderate growth and less pronounced differences.

The lower row reveals the eventual separation between regimes. At $\sigma = 1.50$, output and capital growth rise for several centuries before declining. AI-service growth first slows, then rises again as capital and inference use expand. At year 500, $Y/(AL)$ still grows at 3.2%, above its 2.1% long-run limit, showing a significantly slower transition. A similar pattern of delayed acceleration appears in [Jones \(2026\)](#), drawing on [Jones and Tonetti \(2026\)](#).

I next examine wages and prices in Figure 9. The panels show annual wage growth g_w , the annual net interest rate r , and the level of p_X in final-good units on a logarithmic scale. The

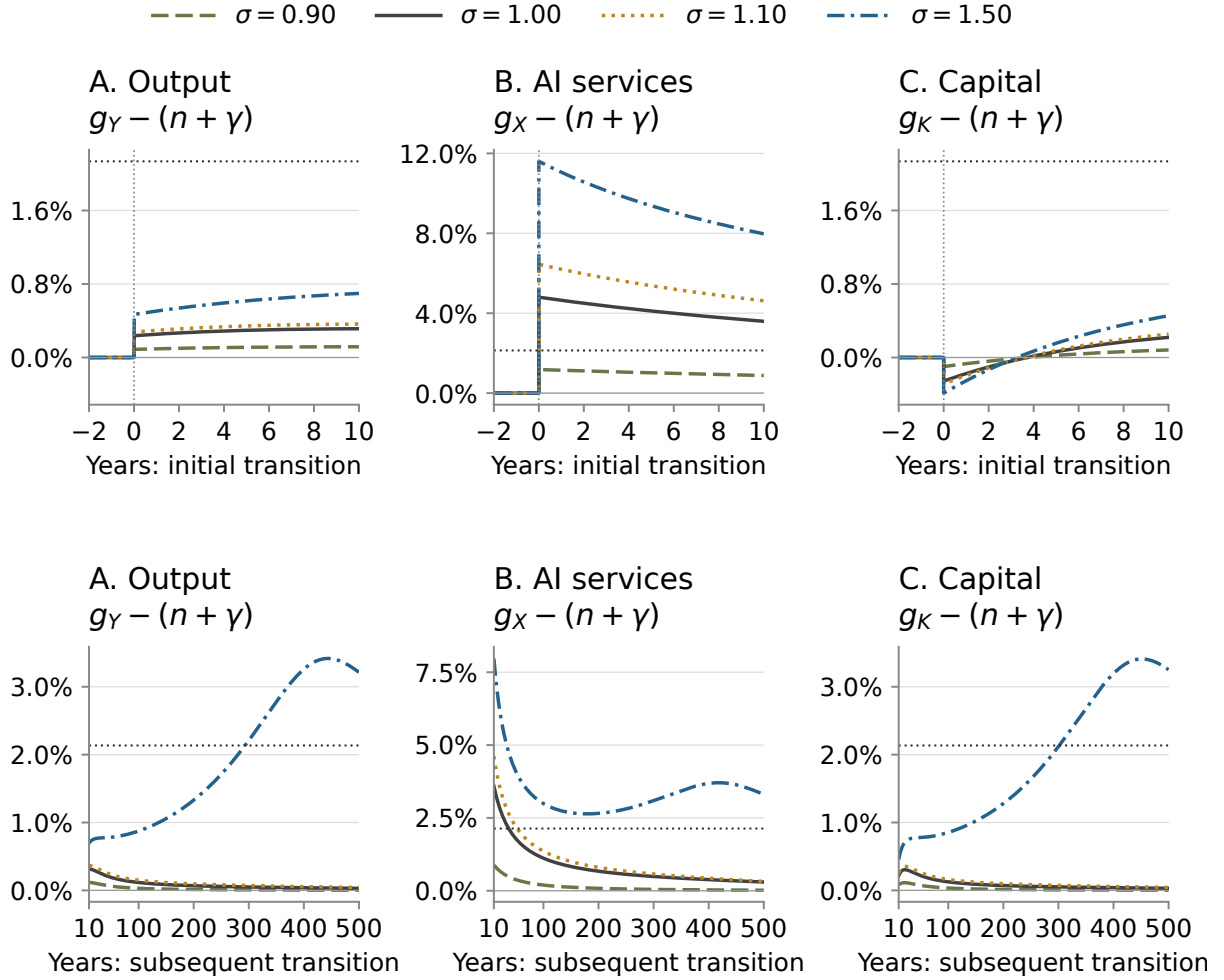


Figure 8: Lower research productivity ($\chi = 1.5$): growth of output, AI production services, and capital per unit of effective labor. The upper row shows two pre-event years and years 0–10; the lower row shows years 10–500. The vertical line marks RSI activation.

top panels show the short run and the bottom panels the long run. Initially, wage growth rises modestly in all four economies. Unlike the preceding exercise, it continues to increase during the first years even in the labor-bottleneck cases: capital accumulation helps sustain the increase in labor's marginal product. The interest rate also rises more gradually as output outpaces capital. At $\sigma = 1.50$, it reaches 5.5% after ten years, compared with 7.8% under higher research productivity. Meanwhile, slower efficiency improvements make the decline in p_X more gradual.

In the bottom row, wage growth and the interest rate eventually diverge between regimes. The three labor-bottleneck cases approach 1.0% and 5.0%, respectively. At $\sigma = 1.50$, the rise and subsequent decline occur much later, and adjustment remains incomplete at year 500: the interest rate is still 8.2%, above its 7.1% long-run value. The limiting wage-growth rate remains 2.4%, below output-per-worker growth of 3.1%. Lower research productivity therefore delays the transition without eliminating the long-run difference between these growth rates.

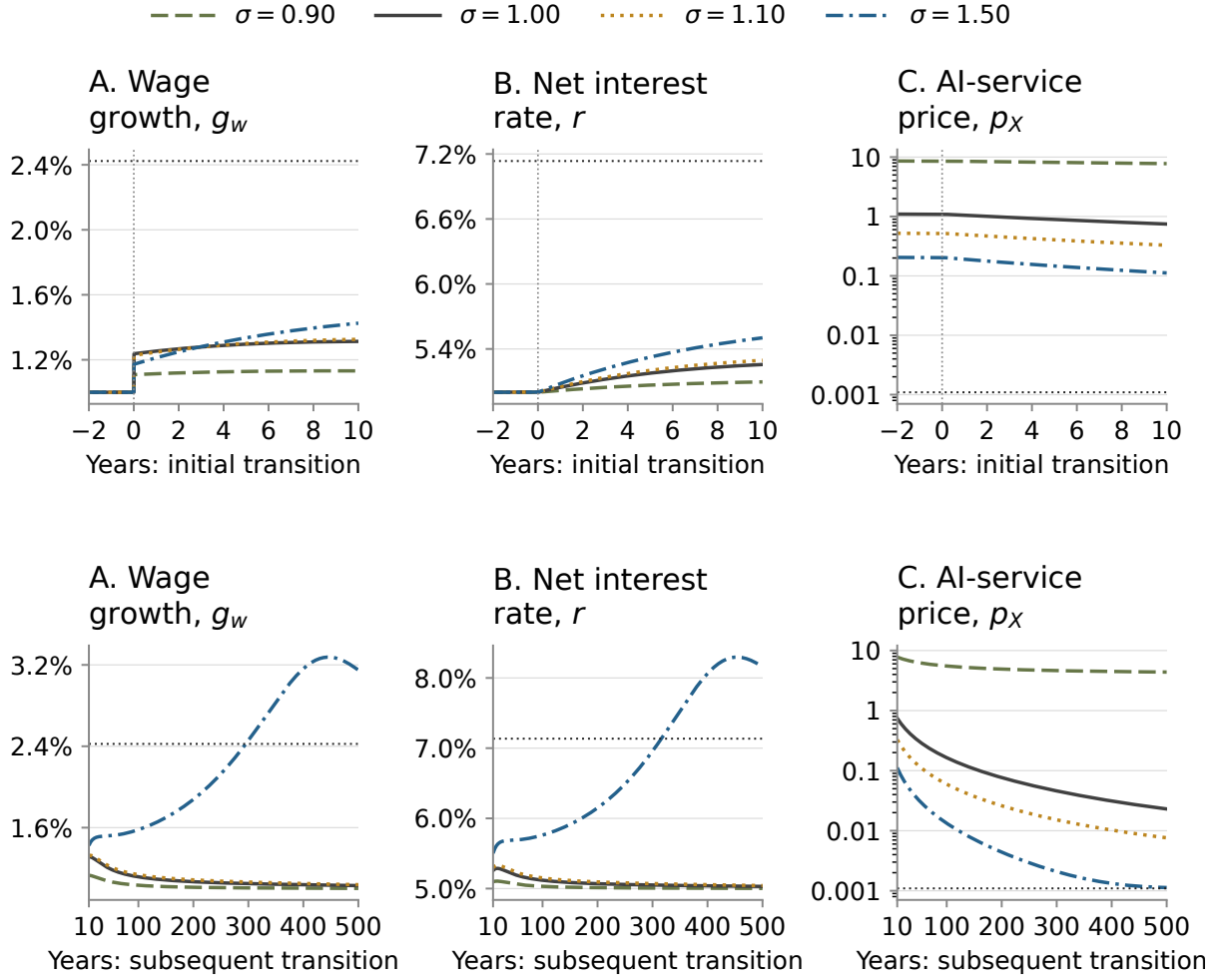


Figure 9: Lower research productivity ($\chi = 1.5$): wage growth, the net interest rate, and the AI-service price. The price is in final-good units on a logarithmic scale. The vertical line marks RSI activation.

I complete the comparison with Figure 10. The panels report wL/Y , Π/Y , U/Y , and M/Y , all as percentages of output. The upper two rows cover the initial transition and the lower two cover years 10–500, with the same scenario lines. The omitted gross capital share remains 33.0%.

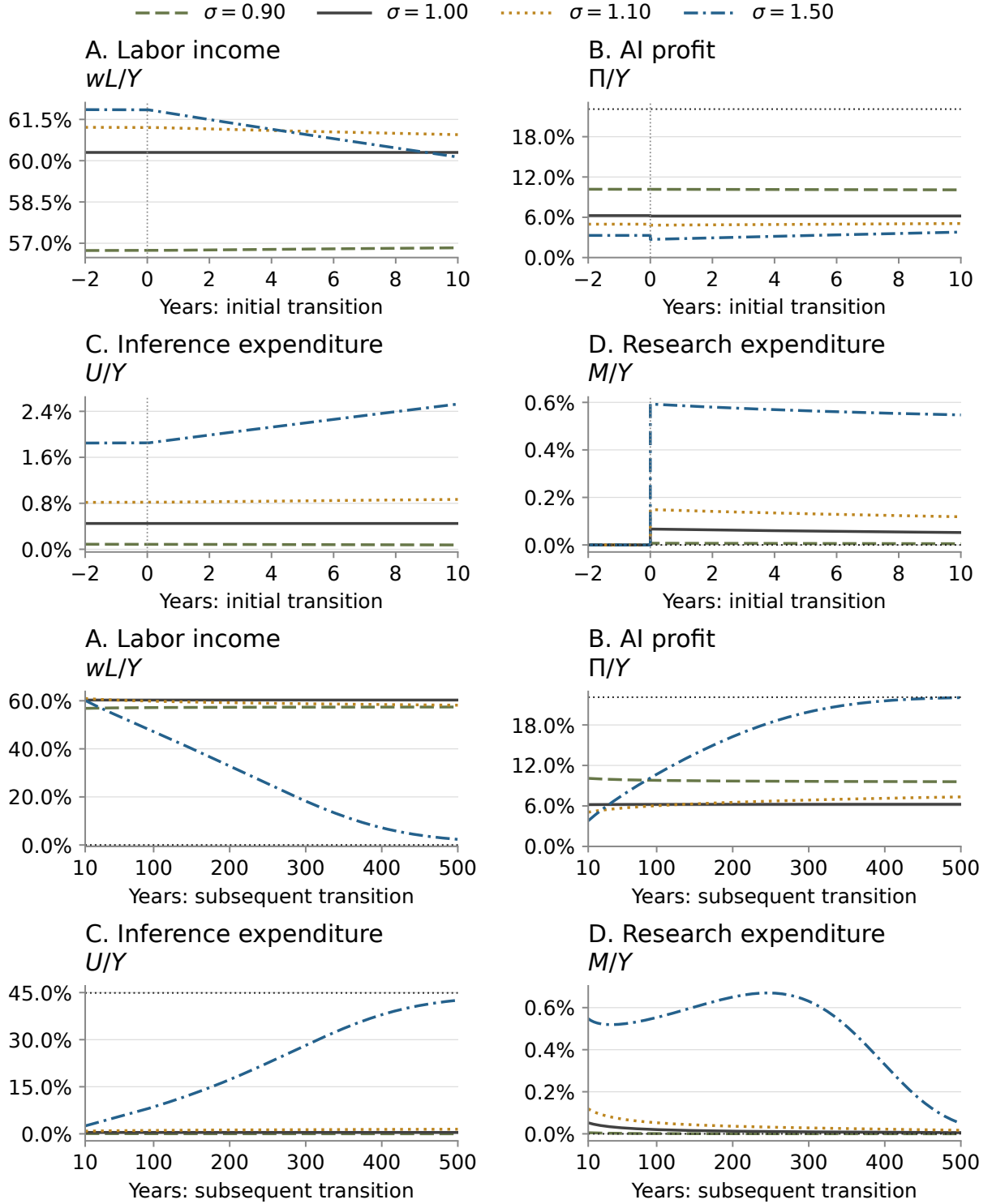


Figure 10: Lower research productivity ($\chi = 1.5$): labor income, AI net profit, inference expenditure, and research expenditure relative to output. The first two rows show the pre-event BGP and years 0–10; the last two show years 10–500. The omitted gross capital share remains $\alpha = 0.33$. Short-run panels A and C use scales fitted to the initial trajectories and omit the long-run reference lines.

In the short run, research absorbs a smaller share of output, leaving more of AI revenue as net profit. At $\sigma = 1.50$, initial research expenditure is 0.6% of output rather than 3.1%, and net profit is 2.7% rather than 0.2%. Research spending reduces net profit at activation, but less sharply than in the preceding exercise. Inference expenditure and labor's income share change little in the labor-bottleneck cases. In the AI-dominated case, inference's share increases and labor's share declines, but both movements are initially gradual, consistent with the smaller output–wage growth gap.

The longer horizon shows the same eventual redistribution as in Figure 7: labor's share tends to zero at $\sigma = 1.50$, while inference expenditure and developer profit approach the same respective limiting shares of output as in the higher-productivity exercise. Research expenditure relative to output first declines, then rises again before falling toward zero. Expanding AI sales provide an incentive for further research, while proximity to the upper bound eventually limits the efficiency gains it can deliver. The other three economies retain a positive labor share. Lower research productivity changes the adjustment and delays the divergence between regimes without changing their limiting growth rates and income shares.

6.4 From competition to monopoly

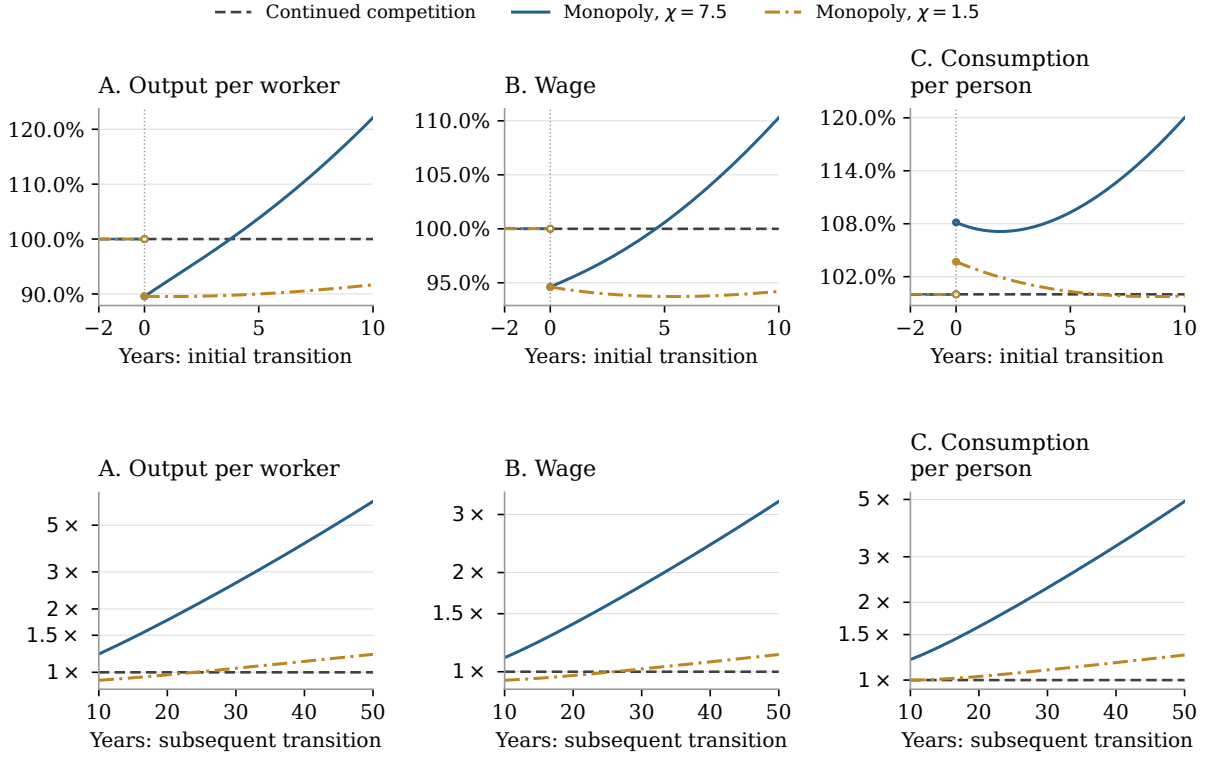
I now ask whether research financed by monopoly rents can compensate for the restriction on AI use caused by monopoly pricing and allow an economy to escape the labor bottleneck. I set $\sigma = 1.5$ and study $\chi = 7.5$ and $\chi = 1.5$. Unlike the preceding exercises, RSI is available before date zero, with the same positive χ as after the change. Research is nevertheless zero under the competitive arrangement in Section 5, because improvements cannot be privately appropriated.

I retain the technology parameters in Table 2, but choose initial capital to place the economy on its competitive balanced-growth path at the same B_0 . Before the change, $p_X = 1/B_0$, $r = \rho + \gamma$, and $K_0/Y_{0-} = 3.3$. Initial efficiency lies below $\mathcal{B}^{\text{mc}}(1.5)$, so continued competition preserves the labor-bottleneck. The upper bound instead exceeds $\mathcal{B}(1.5)$, allowing the monopoly paths considered here to approach AI-dominated growth.

At date zero, an unexpected permanent exclusive right covers both the existing AI technology and its subsequent improvements. In particular, competitive suppliers can no longer use the old technology. This experiment changes research incentives and pricing, not the RSI technology. It abstracts from the cost of enforcing exclusivity and is not an optimal policy exercise.

Figure 11 compares both research productivities with continued competition from the same initial stocks. The three columns show output per worker, wages, and consumption per person, each relative to its counterpart under continued competition.

From competition to monopoly | $\sigma = 1.5$



Top: 100% = competition. Bottom: multiples (log scale), 1x = competition.
Open/filled dots: before/after exclusive rights. Vertical scales differ across windows.

Figure 11: From competition to monopoly with RSI at $\sigma = 1.5$: output per worker, wages, and consumption per person relative to continued competition. The two monopoly paths differ only in χ . Rows cover years $-2-10$ and $10-50$. All three variables use percentages above and multiples on logarithmic scales below; the competitive benchmark is 100% or $1\times$, respectively.

The upper row shows the initial restriction in AI use before efficiency can improve: output falls by 10.4% and wages by 5.4% under either χ . With $\chi = 7.5$, research subsequently raises efficiency enough for output to regain its competitive counterpart around year 3.8 and wages around year 4.6. With $\chi = 1.5$, both remain below their competitive counterparts throughout the first decade. Consumption per person nevertheless rises on impact, by 8.2% with $\chi = 7.5$ and 3.7% with $\chi = 1.5$.

The lower row makes the slower recovery visible: output and wages under $\chi = 1.5$ regain their competitive counterparts around years 24.0 and 26.0. Output and wages subsequently exceed their competitive counterparts under both monopoly paths. Limiting output-per-worker growth, wage growth, and the interest rate are 3.1%, 2.4%, and 7.1% under either χ , as characterized in Proposition 1.

The growth advantage of monopoly in this exercise is not general. As Corollary 1 shows, initial AI efficiency can exceed the competitive threshold while even its upper bound remains below the monopoly threshold. Competition then sustains AI-dominated growth, whereas monopoly-led

research produces more efficient AI but leaves the economy labor-bottlenecked. I illustrate this opposite outcome in the [supplementary simulations available in the digital edition](#).

7 Conclusion

This paper endogenizes investment in AI research, which in turn affects the pace of economic growth. Even under RSI, computing requires resources that compete with consumption and accumulation of physical capital. Algorithms and model weights, by contrast, are non-rival: their use in one computation does not reduce their availability for others. Following [Romer \(1990\)](#), I use monopoly rents to make research investment profitable.

Paradoxically, when AI efficiency is bounded, permanently faster wage growth requires AI to substitute for labor, not complement it. If the upper bound on AI efficiency is sufficiently high, substitution allows capital and AI services to expand faster than effective labor, raising workers' marginal products and sustaining wage growth above the rate of labor-augmenting technological progress. Output per worker grows even faster, so labor's income share vanishes. In the labor-bottleneck regime, by contrast, labor retains a positive income share, but wages and output per worker eventually grow at the rate of labor-augmenting technological progress. Without the upper bound, continued AI efficiency improvements at unit elasticity can sustain a common growth rate for wages and output per worker above the rate of labor-augmenting technological progress, while preserving a constant positive labor share.

The comparison between monopoly and perfect competition in the supply of AI services shows how research incentives and AI pricing jointly determine long-run growth. Under free access to AI models and their improvements, competitive suppliers do not finance RSI. Nevertheless, sufficiently high initial efficiency sustains AI-dominated growth without monopoly or further research. The simulations show that, from an initial efficiency too low to sustain AI-dominated growth under competition, monopoly rents can finance improvements that allow an escape from the labor bottleneck. But markups restrict AI use and raise the efficiency required for AI-dominated growth relative to marginal-cost pricing. Competition can therefore sustain AI-dominated growth where monopoly produces more efficient AI but leaves the economy labor-bottlenecked. Even when both economies are AI-dominated, monopoly delivers faster long-run growth only if its efficiency gains offset the reduction in AI use caused by its markup.

My comparison between monopoly and competition raises the question of how to reward AI research without restricting the use of the technology it produces. Monopoly rents encourage research, but monopoly pricing can leave the economy labor-bottlenecked even when competition would sustain AI-dominated growth. Removing those rents after innovation, however, can undermine incentives to invest beforehand. Patent buyouts offer one possible way to compensate developers while opening access to existing AI technology ([Kremer, 1998](#)). Future research could compare how alternative mechanisms affect investment in AI research, access to the resulting technology, and welfare.

The numerical simulations show that early output-per-worker growth can remain moderate and similar across economies approaching different long-run regimes. The question remains whether human labor will be essential enough for wages to keep pace with output per worker. The answer depends on both technology and the market conditions governing AI development and use.

A Proofs

A.1 Developer optimization

The research problem uses operating profit after AI-service supply has been chosen optimally. I first establish that condition (19) identifies the unique profit-maximizing supply, so I can use this condition when constructing equilibrium trajectories.

Lemma 1 (Optimal AI-service supply). *The developer's operating profit in problem (18) has a unique global maximizer $X > 0$. This optimal supply of AI services is characterized by condition (19) and varies smoothly with K, A, L , and B .*

Proof of Lemma 1. Hold K, A, L, B fixed and let $R(X) = p_X(X)X$ denote AI-service revenue. I show that marginal revenue crosses the constant marginal cost $1/B$ exactly once, from above, so operating profit has a unique global maximum. The inverse-demand elasticity in (15) and differentiation of (10) give

$$R'(X) = p_X(X)(1 - e_X), \quad \frac{ds_X}{d \ln X} = \left(1 - \frac{1}{\sigma}\right) s_X(1 - s_X).$$

Differentiating $R'(X)$ with respect to X gives

$$-\frac{XR''(X)}{p_X(X)} = e_X(1 - e_X) + \left(\alpha - \frac{1}{\sigma}\right) \left(1 - \frac{1}{\sigma}\right) s_X(1 - s_X). \quad (\text{A1})$$

At supply levels with $R'(X) > 0$, we have $0 < e_X < 1$, so $e_X(1 - e_X) > 0$. If $(\alpha - 1/\sigma)(1 - 1/\sigma) \geq 0$, the remaining term is nonnegative, so the right-hand side is positive. If $(\alpha - 1/\sigma)(1 - 1/\sigma) < 0$, the parameter restrictions imply $0 < \alpha < 1/\sigma < 1$. Substituting $e_X = \sigma^{-1}(1 - s_X) + \alpha s_X$ and expanding gives

$$\begin{aligned} -\frac{XR''(X)}{p_X(X)} &= \frac{1}{\sigma} \left(1 - \frac{1}{\sigma}\right) (1 - s_X)^2 \\ &\quad + \left[\frac{1}{\sigma^2} + 2\alpha - \frac{3\alpha}{\sigma}\right] s_X(1 - s_X) + \alpha(1 - \alpha)s_X^2. \end{aligned}$$

All three coefficients in this expansion are positive. For the coefficient of $s_X(1 - s_X)$, this follows from

$$\frac{1}{\sigma^2} + 2\alpha - \frac{3\alpha}{\sigma} = \left(\frac{1}{\sigma} - \alpha\right)^2 + \alpha \left(2 - \frac{1}{\sigma} - \alpha\right) > 0.$$

Hence $R''(X) < 0$ at every supply level where marginal revenue $R'(X)$ is positive.

As $X \downarrow 0$, the price $p_X(X)$ tends to infinity, while $1 - e_X$ converges to a strictly positive value in all three substitution regimes. Hence $R'(X) = p_X(X)(1 - e_X) \rightarrow \infty$. For sufficiently large X , marginal revenue falls below marginal cost $1/B$: it converges to zero from above if $\sigma \geq 1$ and becomes negative if $\sigma < 1$.

Marginal revenue therefore crosses marginal cost at least once. Because $R''(X) < 0$ wherever $R'(X) > 0$, the crossing is unique and occurs from above. The unique solution to $R'(X) = 1/B$ therefore maximizes operating profit globally and satisfies condition (19).

At this solution, $R''(X) < 0$, so the implicit-function theorem applied to $R'(X) = 1/B$ implies that optimal supply is smooth in K, A, L, B for each fixed σ . $\square$

The first-order and transversality conditions alone do not establish that a research trajectory maximizes discounted net profit. The next lemma provides a sufficient condition using only the research elasticity η , the production parameters, and initial AI efficiency relative to its upper bound. For this verification, the given paths of K, A, L are continuous and r is locally integrable.

Lemma 2 (Sufficient conditions for optimal research). *Suppose $\bar{B} < \infty$ and*

$$\eta \left(1 - \frac{B_0}{\bar{B}} \right) \leq \min \left\{ \alpha, \frac{1}{\sigma} \right\}. \quad (\text{A2})$$

Consider a feasible research trajectory $\{B_t, M_t\}_{t \geq 0}$ for problem (21), with $M_t > 0$ at every date and a finite present value of net profit. If there exists a continuously differentiable shadow-price trajectory $\{q_t\}_{t \geq 0}$ satisfying the research condition (22), the costate equation (23), and the transversality condition (25), then the research trajectory maximizes the developer's discounted net profit.

Proof of Lemma 2. I first bound the response of optimal AI-service supply to efficiency. I then use this bound and condition (A2) to make the research-maximized Hamiltonian concave after a change of state variable. Concavity bounds the gain from any alternative trajectory by a quantity that vanishes under the candidate's transversality condition.

1. *Bounding the response of AI-service supply.* Let $X_t^*(b)$ denote optimal supply at efficiency b , and let $\pi_t(b)$ denote the resulting operating profit in problem (18). Lemma 1 ensures that this supply is unique and smooth. Implicit differentiation of the monopoly condition (19) gives its elasticity with respect to b :

$$\varepsilon_{B,t}(b) \equiv \frac{\partial \ln X_t^*(b)}{\partial \ln b} = \frac{1 - e_X}{e_X(1 - e_X) + (\alpha - \sigma^{-1})(1 - \sigma^{-1})s_X(1 - s_X)}. \quad (\text{A3})$$

Here s_X and e_X are evaluated at $X_t^*(b)$. The denominator is positive by equation (A1), and $0 < e_X < 1$ at the optimum. If $(\alpha - 1/\sigma)(1 - 1/\sigma) \geq 0$, the denominator is at least $e_X(1 - e_X)$, with $e_X \geq \min\{\alpha, 1/\sigma\}$. In the remaining case, $\alpha < 1/\sigma < 1$. Substituting $e_X = \sigma^{-1}(1 - s_X) + \alpha s_X$ and subtracting $\alpha(1 - e_X)$ from the denominator gives

$$(\sigma^{-1} - \alpha)(1 - s_X) [(1 - \sigma^{-1})(1 - s_X) + (\sigma^{-1} - \alpha)s_X] \geq 0.$$

Thus, in both cases, the denominator is at least $\min\{\alpha, 1/\sigma\}(1 - e_X)$, so

$$0 < \varepsilon_{B,t}(b) \leq \max\{\sigma, 1/\alpha\}. \quad (\text{A4})$$

This bound holds at every date and every positive efficiency, independently of the quantities of capital and effective labor.

2. *Transforming AI efficiency and maximizing over research.* Every feasible trajectory remains below $\bar{B}$ at finite dates. Indeed, feasibility requires nonnegative, locally integrable research expenditure and a locally absolutely continuous state satisfying the RSI technology (17) almost everywhere. Integrating that law gives

$$\bar{B} - B_t = (\bar{B} - B_0) \exp \left\{ -\frac{\chi}{\bar{B}} \int_0^t (B_s M_s)^\eta ds \right\} > 0. \quad (\text{A5})$$

Continuity of B , local integrability of M , and $0 < \eta < 1$ make the exponent finite at every finite date. Hence the relevant efficiency domain, for both the candidate and every feasible alternative, is $[B_0, \bar{B})$.

I express the research problem in terms of an increasing transformation $S(B)$ of AI efficiency. The transformation is chosen so that condition (A2) ensures concavity of the maximized Hamiltonian, as verified in Step 3. Define

$$v = \frac{1 - \eta}{\eta \psi(B_0)}, \quad p = \frac{1}{\eta \psi(B_0)} - 1, \quad S(B) = \int_0^B \frac{z^{v-1}}{\psi(z)} dz. \quad (\text{A6})$$

Both exponents are positive. The transformation maps $[B_0, \bar{B})$ increasingly onto $[S(B_0), \infty)$ and has inverse $B(S)$, with derivative $B'(S) = \psi(B)B^{1-v}$.

Writing $S_t = S(B_t)$, the chain rule applied to the RSI technology gives

$$\dot{S}_t = \chi B_t^{v-1+\eta} M_t^\eta.$$

The shadow price in the new coordinates is $q_{S,t} = q_t B'(S_t) = q_t \psi(B_t) B_t^{1-v} > 0$.

At each date, hold this shadow price at its candidate value and maximize the transformed Hamiltonian over research expenditure:

$$\hat{\mathcal{H}}_t(S) = \max_{\widetilde{M} \geq 0} \left\{ \pi_t(B(S)) - \widetilde{M} + q_{S,t} \chi B(S)^{v-1+\eta} \widetilde{M}^\eta \right\}.$$

The expression is strictly concave in $\widetilde{M}$. Its first-order condition, together with the candidate's research condition (22), gives $\widetilde{M} = M_t [B(S)/B_t]^p$, since $p = (v - 1 + \eta)/(1 - \eta)$. Substituting this choice yields

$$\hat{\mathcal{H}}_t(S) = \pi_t(B(S)) + \frac{1 - \eta}{\eta} M_t \left[\frac{B(S)}{B_t} \right]^p. \quad (\text{A7})$$

3. *Establishing concavity.* Fix a date and write $b = B(S)$. I show that each term in (A7) is concave

in S over the entire feasible efficiency domain. The envelope condition (20) gives $\pi'_t(b) = X_t^*(b)/b^2$. Differentiating this expression gives $\pi''_t(b) = X_t^*(b)[\varepsilon_{B,t}(b) - 2]/b^3$. Using $B'(S) = \psi(b)b^{1-v}$ and $\psi'(b) = -1/\bar{B}$, the chain rule therefore gives

$$\frac{d^2\pi_t(B(S))}{dS^2} = X_t^*(b)b^{-1-2v}\psi(b) \left[(\varepsilon_{B,t}(b) - 1 - v)\psi(b) - \frac{b}{\bar{B}} \right].$$

Applying the same chain rule to the power $B(S)^p$ in the research term gives

$$\frac{d^2B(S)^p}{dS^2} = p b^{p-2v}\psi(b) \left[(p - v)\psi(b) - \frac{b}{\bar{B}} \right].$$

The parameter condition (A2) and the supply-elasticity bound (A4) imply

$$\varepsilon_{B,t}(b) \leq \frac{1}{\eta\psi(B_0)} = p + 1.$$

The bracket in the operating-profit derivative is therefore no greater than the bracket in the research-term derivative. Substitution of p , v , and $\psi(b) = 1 - b/\bar{B}$ reduces the latter bracket to

$$(p - v)\psi(b) - \frac{b}{\bar{B}} = \frac{B_0 - b}{\bar{B} - B_0} \leq 0.$$

Both second derivatives are nonpositive for every $b \geq B_0$. Consequently, the maximized Hamiltonian is concave in S and lies below its tangent at the candidate state S_t .

4. *Comparing discounted net profits.* Let $D_t = \exp(-\int_0^t r_s ds)$ and $\mu_t = D_t q_{S,t}$. Differentiating (A7) at S_t and using the envelope condition gives

$$\hat{\mathcal{H}}'_t(S_t) = \psi(B_t)B_t^{-v} \left[U_t + \frac{1 - \eta}{\eta} p M_t \right].$$

The research condition and the RSI law imply $q_t \eta \dot{B}_t = M_t$. Differentiating $q_{S,t} = q_t \psi(B_t)B_t^{1-v}$ and using this identity and the costate equation (23) gives

$$\dot{q}_{S,t} = r_t q_{S,t} - \hat{\mathcal{H}}'_t(S_t), \quad \dot{\mu}_t = -D_t \hat{\mathcal{H}}'_t(S_t).$$

Consider any feasible alternative $\{\tilde{B}_t, \tilde{M}_t\}_{t \geq 0}$ starting from the same B_0 , and write $\tilde{S}_t = S(\tilde{B}_t)$. At the candidate shadow price, the alternative's Hamiltonian is at most $\hat{\mathcal{H}}_t(\tilde{S}_t)$, whereas the candidate attains $\hat{\mathcal{H}}_t(S_t)$. The Hamiltonian equals current net profit plus $q_{S,t}\dot{S}_t$. Subtracting these state-accumulation terms and applying concavity therefore bounds the difference in discounted net

profit, almost everywhere:

$$\begin{aligned}
& D_t \left[\pi_t(\tilde{B}_t) - \tilde{M}_t - \pi_t(B_t) + M_t \right] \\
& \leq D_t \hat{\mathcal{H}}'_t(S_t)(\tilde{S}_t - S_t) - \mu_t(\dot{\tilde{S}}_t - \dot{S}_t) \\
& = -\frac{d}{dt} \left[\mu_t(\tilde{S}_t - S_t) \right].
\end{aligned}$$

The last equality uses the transformed costate equation.

Let J_T and $\tilde{J}_T$ denote the candidate's and the alternative's discounted net profits through date T . Integrating and using their common initial state gives

$$\tilde{J}_T - J_T \leq -\mu_T(\tilde{S}_T - S_T) \leq \mu_T S_T,$$

because $\mu_T > 0$ and $\tilde{S}_T \geq 0$. Thus any gain from the alternative is bounded above by the candidate's terminal value $\mu_T S_T$.

It remains to show that this terminal value vanishes. Since $\psi(z) \geq \psi(B_t)$ for $0 \leq z \leq B_t$, the definition of S gives

$$0 \leq q_{S,t} S_t = q_t \psi(B_t) B_t^{1-v} \int_0^{B_t} \frac{z^{v-1}}{\psi(z)} dz \leq \frac{q_t B_t}{v}.$$

The candidate's transversality condition therefore implies $\mu_T S_T \leq D_T q_T B_T / v \rightarrow 0$. Because its present value of net profit is finite,

$$\limsup_{T \rightarrow \infty} \tilde{J}_T \leq \lim_{T \rightarrow \infty} J_T < \infty.$$

No feasible alternative yields a higher present value of net profit. □

Condition (A2) is sufficient, not necessary. For any fixed production parameters and research elasticity η , it holds when B_0 is sufficiently close to $\bar{B}$. The lemma verifies optimality of a candidate trajectory; it does not establish that such a trajectory exists.

A.2 Equilibria with bounded AI efficiency

Proof of Proposition 1. I first derive candidate long-run allocations from the static optimum in Lemma 1. I then normalize the dynamic equations and use the stable-manifold and inverse-function theorems to construct convergent trajectories from nearby initial stocks. Finally, I verify feasibility, market clearing, both transversality conditions, and global agent optimality. The result is local in initial stocks, but the constructed equilibria exist for all future time.

Step 1. Feasible efficiency and candidate long-run values.

Equation (A5) ensures $B_t < \bar{B}$ at every finite date. I impose $B = \bar{B}$ only in the normalized boundary equations used to identify candidate limits.

Write $\theta = (1 - \alpha)/\alpha$. For $\sigma \neq 1$, let $\varphi = (\sigma - 1)/\sigma$ and define

$$m(s) = 1 - \frac{1-s}{\sigma} - \alpha s, \quad \mathcal{H}_\sigma(s) = \left[(1 - \alpha)s m(s) \left(\frac{\omega_X}{s} \right)^{1/\varphi} \right]^\theta.$$

Here $m(s) = 1 - e_X$ is the marginal-revenue factor, not research compute. The static monopoly optimum has $m(s_X) > 0$ by Lemma 1. The CES share, monopoly condition, and production function, equations (27d), (27i), and (27c), imply

$$\frac{X}{AN} = \left[\frac{\omega_L s_X}{\omega_X (1 - s_X)} \right]^{1/\varphi}, \quad \frac{U}{Y} = (1 - \alpha)s_X m(s_X), \quad (\text{A8})$$

$$\frac{Y}{K} = B^\theta \mathcal{H}_\sigma(s_X). \quad (\text{A9})$$

The identities $Z = X(\omega_X/s_X)^{1/\varphi}$ and $X = B(1 - \alpha)s_X m(s_X)Y$ give the output–capital ratio in (A9).

If $0 < \sigma < 1$, positive marginal revenue requires

$$s > s_{\min} \equiv \frac{1 - \sigma}{1 - \alpha\sigma}.$$

On $(s_{\min}, 1)$, both $m(s)$ and $s^{-1/(\sigma-1)}$ increase strictly with s . Since

$$\mathcal{H}_\sigma(s)^{1/\theta} = (1 - \alpha)\omega_X^{1/\varphi} m(s)s^{-1/(\sigma-1)},$$

the share map $\mathcal{H}_\sigma(s)$ increases from zero to $\mathcal{H}_\sigma(1) = [(1 - \alpha)^2 \omega_X^{1/\varphi}]^\theta$. Setting $r = \rho + \gamma$ in (11) and using (A9) gives the candidate terminal equation:

$$\frac{\rho + \gamma + \delta}{\alpha} = \bar{B}^\theta \mathcal{H}_\sigma(\bar{s}_X) \quad (\text{A10})$$

This equation has exactly one interior solution if and only if $\bar{B} > \mathcal{B}(\sigma)$.

If $\sigma > 1$, $\mathcal{H}_\sigma(s)$ decreases strictly with s on $(0, 1)$. To verify the sign, put $a = \sigma^{-1} - \alpha$, so $m(s) = \varphi + as$. The logarithmic derivative of $\mathcal{H}_\sigma(s)$ with respect to s has the sign of $a(\sigma - 2)s - \varphi$. If $a(\sigma - 2) \leq 0$, then $a(\sigma - 2)s - \varphi < 0$. If $\sigma > 2$ and $a > 0$, $a(\sigma - 2) < (\sigma - 2)/\sigma < \varphi$; if $1 < \sigma < 2$ and $a < 0$, $a(\sigma - 2) < (1 - \sigma^{-1})(2 - \sigma) < \varphi$. The map therefore decreases from infinity to $\mathcal{H}_\sigma(1)$. Equation (A10) has exactly one interior solution if and only if $\bar{B} < \mathcal{B}(\sigma)$. Since $\mathcal{R}(B) = \alpha B^\theta \mathcal{H}_\sigma(1) - \delta$, these calculations also derive (34)–(35). The interior-share condition is therefore $\mathcal{R}(\bar{B}) > \rho + \gamma$ for $\sigma < 1$, but $\mathcal{R}(\bar{B}) < \rho + \gamma$ for $\sigma > 1$.

For either interior-share case, let x, y, u, k, c denote X, Y, U, K, C divided by AN , respectively.

Construct their candidate limiting values sequentially:

$$\begin{aligned}
\bar{x} &= \left[\frac{\omega_L \bar{s}_X}{\omega_X (1 - \bar{s}_X)} \right]^{1/\varphi}, & \bar{u} &= \frac{\bar{x}}{\bar{B}}, \\
\bar{y} &= \frac{\bar{u}}{(1 - \alpha) \bar{s}_X m(\bar{s}_X)}, & \bar{k} &= \frac{\alpha \bar{y}}{\rho + \gamma + \delta}, \\
\bar{c} &= \bar{y} - \bar{u} - (\delta + n + \gamma) \bar{k}, & \bar{r} &= \rho + \gamma.
\end{aligned} \tag{A11}$$

At $\sigma = 1$, define $\beta = (1 - \alpha)\omega_X$. Combining the compute-share identity in (A8) with the unit-elasticity limit of (7)–(8) gives

$$y = \left[(\beta^2 B)^\beta k^\alpha \right]^{1/(1-\beta)}, \quad u = \beta^2 y, \quad x = Bu.$$

The equation $y/k = (\rho + \gamma + \delta)/\alpha$ has exactly one positive solution for every finite $\bar{B} > 0$:

$$\bar{k} = \left[(\beta^2 \bar{B})^\beta \left(\frac{\rho + \gamma + \delta}{\alpha} \right)^{\beta-1} \right]^{1/(1-\alpha-\beta)}. \tag{A12}$$

Set $\bar{y} = [(\rho + \gamma + \delta)/\alpha] \bar{k}$, $\bar{u} = \beta^2 \bar{y}$, and $\bar{x} = \bar{B} \bar{u}$, with $\bar{c}, \bar{r}$ as in (A11).

Consumption positivity does not require an additional restriction. Positive marginal revenue implies $0 < m(\bar{s}_X) < 1$. Since $0 < \bar{s}_X < 1$, in all three cases

$$\frac{\bar{u}}{\bar{y}} = (1 - \alpha) \bar{s}_X m(\bar{s}_X) < 1 - \alpha, \quad \frac{(\delta + n + \gamma) \bar{k}}{\bar{y}} = \frac{\alpha(\delta + n + \gamma)}{\rho + \gamma + \delta} < \alpha.$$

It follows that $\bar{c}/\bar{y} > 0$.

Step 2. The three positive-labor-share cases.

The cases in Proposition 1(i), 1(ii), and 1(iii) share one normalization. Divide capital, consumption, and the AI-efficiency shadow price by effective labor, and multiply the remaining efficiency gap by effective labor:

$$k = \frac{K}{AN}, \quad c = \frac{C}{AN}, \quad d = (\bar{B} - B)AN, \quad \tilde{q} = \frac{q}{AN}, \quad \tau = (AN)^{-1}. \tag{A13}$$

Dividing by AN removes effective-labor growth. Multiplying the shrinking gap by AN allows the scaled gap d to converge to a positive value: $\dot{d}/d = n + \gamma - \dot{B}/(\bar{B} - B)$. The variable d measures the scaled gap, not its rate of closure. The constants $\bar{k}, \bar{d}$ are the limiting values of k, d ; they are determined by the model, not additional parameters. Including τ as a state removes explicit calendar time from the normalized equations. Since $\dot{\tau} = -(n + \gamma)\tau$, its limit is zero. These coordinates leave the model unchanged: $K = k/\tau$ and $B = \bar{B} - d\tau$ at every finite date.

Given k, B , solve the unique static monopoly problem for y, x, u , and set $r = \alpha y/k - \delta$. Define $a = \dot{B}/(\bar{B} - B)$, the proportional rate of closure of the AI-efficiency gap. The research condition

gives

$$B = \bar{B} - d\tau, \quad M = \left[\tilde{q} d \chi \eta \frac{B^\eta}{\bar{B}} \right]^{1/(1-\eta)}, \quad a = \chi \frac{B^\eta M^\eta}{\bar{B}}.$$

Substituting (27k) and (28) gives the exact dynamic equations

$$\frac{\dot{k}}{k} = \frac{y - c - u - M\tau}{k} - \delta - n - \gamma, \quad (\text{A14a})$$

$$\frac{\dot{c}}{c} = r - \rho - \gamma, \quad (\text{A14b})$$

$$\frac{\dot{d}}{d} = n + \gamma - a, \quad (\text{A14c})$$

$$\frac{\dot{\tilde{q}}}{\tilde{q}} = r - \frac{x}{\tilde{q} B^2} - \eta \frac{a d \tau}{B} + a - n - \gamma, \quad (\text{A14d})$$

$$\dot{\tau} = -(n + \gamma)\tau. \quad (\text{A14e})$$

At $\tau = 0$, the stationarity conditions for d and $\tilde{q}$ in (A14), together with the research relations above, give

$$\bar{M} = \left[\frac{(n + \gamma) \bar{B}^{1-\eta}}{\chi} \right]^{1/\eta}, \quad \bar{\tilde{q}} = \frac{\bar{x}}{\bar{r} \bar{B}^2}, \quad \bar{d} = \frac{\bar{M}}{\eta(n + \gamma) \bar{\tilde{q}}}. \quad (\text{A15})$$

All are positive. Direct substitution shows that $(\bar{k}, \bar{c}, \bar{d}, \bar{\tilde{q}}, 0)$ is a fixed point of (A14). This boundary point is a device for constructing paths with $\tau > 0$ and $B < \bar{B}$, not an admissible date-zero state itself.

I next verify saddle-path stability to determine whether initial consumption and the shadow price can be selected so that nearby initial stocks lead to convergence.

Write $\mu = \eta/(1 - \eta)$, $r_k = \partial r / \partial \log k$, and $a_k = \partial[(y - u)/k] / \partial \log k + \bar{c}/\bar{k}$, evaluated at the fixed point. Differentiating (A14) in $(\log k, \log c, \log d, \log \tilde{q}, \tau)$ coordinates gives two diagonal Jacobian blocks: one for capital and consumption, and one for the scaled efficiency gap and normalized shadow price,

$$J_R = \begin{pmatrix} a_k & -\bar{c}/\bar{k} \\ r_k & 0 \end{pmatrix}, \quad J_B = \begin{pmatrix} -(n + \gamma)\mu & -(n + \gamma)\mu \\ (n + \gamma)\mu & \bar{r} + (n + \gamma)\mu \end{pmatrix}, \quad (\text{A16})$$

and the additional eigenvalue $-(n + \gamma)$. In the linearized system, deviations in τ can affect both blocks, and deviations in (k, c) can affect $(d, \tilde{q})$, but not conversely.

To check $r_k < 0$, let $D(s_X)$ be the positive curvature expression in (A1). Implicit differentiation of the monopoly condition gives

$$\frac{\partial \log X}{\partial \log K} = \frac{\alpha(1 - e_X)}{D(s_X)}, \quad r_k = (r + \delta)(1 - \alpha) \left[\frac{\alpha s_X(1 - e_X)}{D(s_X)} - 1 \right].$$

Moreover,

$$D(s_X) - \alpha s_X(1 - e_X) = (1 - s_X) \left\{ \frac{1 - e_X}{\sigma} + \left(\alpha - \frac{1}{\sigma} \right) \left(1 - \frac{1}{\sigma} \right) s_X \right\} > 0.$$

The expression in braces is positive when $1/\sigma \geq 1$ or $1/\sigma \leq \alpha$. For $\alpha < 1/\sigma < 1$, the expression is linear in s_X , with positive values $(1/\sigma)(1 - 1/\sigma)$ at $s_X = 0$ and $1/\sigma^2 + \alpha(1 - 2/\sigma)$ at $s_X = 1$. The value at $s_X = 1$ is positive if $1/\sigma \leq 1/2$; otherwise $\alpha < 1/\sigma$ implies it exceeds $(1/\sigma)(1 - 1/\sigma) > 0$. Thus $r_k < 0$ in every interior-share case.

Both blocks have negative determinants:

$$\det J_R = (\bar{c}/\bar{k})r_k < 0, \quad \det J_B = -(n + \gamma)\mu\bar{r} < 0.$$

There are therefore three stable and two unstable eigenvalues. The stable direction in J_R has a nonzero k component, and the stable direction in J_B has a nonzero d component. Counting stable directions is not enough: their projection must also allow the three predetermined coordinates to vary independently. To verify this, consider a vector in the stable subspace whose state components (k, d, τ) are all zero. Its linearized τ path is zero. With the τ component equal to zero, the capital-consumption block must follow its stable direction; its zero initial k component forces its c component to be zero as well. With the capital and consumption components also zero, the same argument in the research block makes the $\tilde{q}$ component zero. The projection of the three-dimensional stable subspace onto the three predetermined coordinates is therefore injective and hence invertible. This argument also covers repeated stable eigenvalues.

Lemma 1 makes the static solution smooth. The right-hand sides of the normalized dynamic equations extend smoothly through $\tau = 0$, since $B = \bar{B} - d\tau$ remains positive nearby. The stable-manifold theorem therefore gives a local manifold whose trajectories converge to the fixed point. The invertible projection just established lets the inverse-function theorem express this manifold as a local graph selecting $(c, \tilde{q})$ from (k, d, τ) ; see Dyatlov (2018).

Choose a sufficiently small neighborhood of the full fixed point and then a smaller initial-state neighborhood $\mathcal{U}_L$ of $(\bar{k}, \bar{d}, 0)$. These neighborhoods may depend on the parameters and the case under consideration. In original initial stocks, the restriction is

$$(k_0, d_0, \tau_0) = \left(\frac{K_0}{A_0 N_0}, (\bar{B} - B_0)A_0 N_0, \frac{1}{A_0 N_0} \right) \in \mathcal{U}_L. \quad (\text{A17})$$

Closeness is required in all three coordinates; B_0 close to $\bar{B}$ alone is not sufficient. For every physical initial state satisfying (A17), the graph selects a solution that stays in the full neighborhood for all future time and converges to the fixed point. It is unique among solutions with those properties.

The full neighborhood can be chosen so that $k, d, c, \tilde{q}$ and $B = \bar{B} - d\tau$ remain positive. For a physical initial state, $\tau_t > 0$ at every finite date, so $B_t < \bar{B}$; the research condition also gives $M_t > 0$. Reconstruction recovers the original initial stocks and equilibrium equations. The right-hand sides of the normalized dynamic equations tend to zero, so the smooth ratios $K/(AN)$, $C/(AN)$, $Y/(AN)$, and w/A have time derivatives tending to zero. These ratios also converge to positive constants, implying $g_K, g_C, g_Y \rightarrow n + \gamma$ and $g_w \rightarrow \gamma$. Together with the candidate limits of r and s_X , this establishes (36)–(38). Optimality and transversality are verified below.

Step 3. *The AI-dominated case.*

Let $\sigma > 1$ and $\bar{B} > \mathcal{B}(\sigma)$. Here $\varphi = (\sigma - 1)/\sigma \in (0, 1)$. I seek an equilibrium in which capital grows faster than effective labor. Since $K/(AN)$ then diverges, I use capital rather than effective labor to normalize consumption, the efficiency gap, and the shadow price. The two sets of coordinates are:

	Labor bottleneck (Step 2)	AI-dominated (Step 3)
Relative stocks	$k = K/(AN)$	$h = (AN/K)^\varphi$
Consumption	$c = C/(AN)$	$c = C/K$
Efficiency gap	$d = (\bar{B} - B)AN$	$d = (\bar{B} - B)K$
Shadow price	$\tilde{q} = q/(AN)$	$v = q/K$

In both cases $\tau = (AN)^{-1}$. The exponent φ makes h enter the CES sum linearly, so the zero-labor-share limit can be studied at $h = 0$. Within this construction, y, x, u denote $Y/K, X/K, U/K$, respectively. The constant $\bar{d}$ now denotes the limit of $(\bar{B} - B)K$, not of $(\bar{B} - B)AN$; no original state or control is redefined. Rewriting the relations used in Step 2 in these coordinates gives

$$B = \bar{B} - d\tau h^{1/\varphi}, \quad M = [vd\chi\eta B^\eta/\bar{B}]^{1/(1-\eta)}, \quad a = \chi B^\eta M^\eta/\bar{B}, \quad r = \alpha y - \delta.$$

Using (28), capital's growth rate is $g_K = y - c - u - M\tau h^{1/\varphi} - \delta$. Differentiating the normalized variables then gives

$$\dot{h} = \varphi h(n + \gamma - g_K), \tag{A18a}$$

$$\dot{d}/d = g_K - a, \tag{A18b}$$

$$\dot{c}/c = n + r - \rho - g_K, \tag{A18c}$$

$$\dot{v}/v = r - x/(vB^2) - \eta ad\tau h^{1/\varphi}/B + a - g_K, \tag{A18d}$$

$$\dot{\tau} = -(n + \gamma)\tau. \tag{A18e}$$

The CES term divided by capital is $[\omega_L h + \omega_X x^\varphi]^{1/\varphi}$. At $h = 0$ the monopoly equation has the strictly positive curvature $D(1) = \alpha(1 - \alpha)$. The static monopoly solution is therefore smooth in (h, B) near $(0, \bar{B})$. To apply the stable-manifold theorem at $h = 0$, I extend the right-hand sides of the normalized dynamic equations to an open neighborhood of this boundary. For the remaining terms, replacing $h^{1/\varphi}$ by $\max\{h, 0\}^{1/\varphi}$ outside the physical domain gives a C^1 extension because $1/\varphi > 1$. This meets the regularity requirement of the C^1 stable-manifold theorem; see [Feldman \(2021, p. 7\)](#). The extension is only a mathematical device and does not change any equation on $h \geq 0$.

The limiting interest and aggregate growth rates are

$$\bar{r} = \mathcal{R}(\bar{B}), \quad \bar{g} = n + \bar{r} - \rho > n + \gamma. \tag{A19}$$

At $h = \tau = 0$, combining the static limits from Step 1 with the stationarity conditions in (A18) gives

$$\begin{aligned}\bar{y} &= \frac{\mathcal{R}(\bar{B}) + \delta}{\alpha}, & \bar{u} &= (1 - \alpha)^2 \bar{y}, & \bar{x} &= \bar{B} \bar{u}, \\ \bar{r} &= \mathcal{R}(\bar{B}), & \bar{g} &= n + \bar{r} - \rho, & \bar{c} &= \alpha(1 - \alpha)\bar{y} + \rho - n > 0, \\ \bar{v} &= \bar{x}/(\bar{r} \bar{B}^2), & \bar{M} &= [\bar{g} \bar{B}^{1-\eta}/\chi]^{1/\eta}, & \bar{d} &= \bar{M}/(\eta \bar{g} \bar{v}).\end{aligned}\tag{A20}$$

The threshold implies $\bar{g} > n + \gamma > 0$ and $\bar{r} > \rho + \gamma > 0$, so all constants are positive. Substitution verifies the fixed point $(0, \bar{d}, \bar{c}, \bar{v}, 0)$.

The h and τ eigenvalues are $\varphi(n + \gamma - \bar{g}) < 0$ and $-(n + \gamma) < 0$. The first equals $\varphi[\rho + \gamma - \mathcal{R}(\bar{B})]$. It becomes zero at $\mathcal{R}(\bar{B}) = \rho + \gamma$, which is why this stable-manifold argument does not cover the threshold. The consumption-ratio eigenvalue is $\bar{c} > 0$. The remaining block in $(\log d, \log v)$ is

$$\begin{pmatrix} -\bar{g}\mu & -\bar{g}\mu \\ \bar{g}\mu & \bar{r} + \bar{g}\mu \end{pmatrix}, \quad \det = -\bar{g}\mu\bar{r} < 0.$$

The linearized system is triangular in the order $(h, \tau), c, (d, v)$. A stable vector with zero state components (h, d, τ) has zero h, τ paths. With those components zero, the linearized consumption equation has a positive coefficient $\bar{c}$; its consumption component must therefore also be zero for convergence. The research block then has zero d component and hence zero v component, as in the previous construction. Thus the three-dimensional stable subspace projects invertibly onto (h, d, τ) . The same graph argument gives a neighborhood $\mathcal{U}_A$ of $(0, \bar{d}, 0)$ such that the initial-state condition is

$$\left[\left(\frac{A_0 N_0}{K_0} \right)^\varphi, (\bar{B} - B_0)K_0, \frac{1}{A_0 N_0} \right] \in \mathcal{U}_A.\tag{A21}$$

The graph determines a convergent trajectory for every admissible initial state in this neighborhood, with a unique nearby convergent choice of C_0, q_0 and positive quantities at every finite date.

The right-hand sides of the normalized dynamic equations tend to zero, so the time derivatives of the smooth positive ratios Y/K , X/K , and C/K tend to zero. Consequently $g_Y, g_X, g_C, g_K \rightarrow \bar{g}$ and $s_X \rightarrow 1$. Log differentiation of the CES share in equation (27d) and the wage equation (27g) gives the exact identity

$$g_w = g_Y - n + \frac{\sigma - 1}{\sigma} s_X (n + \gamma - g_X).$$

Taking limits yields

$$g_w \rightarrow \bar{g} - n + \frac{\sigma - 1}{\sigma} (n + \gamma - \bar{g}) = \gamma + \frac{\bar{r} - \rho - \gamma}{\sigma}.$$

This proves the growth and distribution limits in Proposition 1(iv).

Step 4. Optimality and transversality.

In either construction let $\bar{g}$ denote the limiting aggregate growth rate: $n + \gamma$ in the positive

labor-share case and $n + \bar{r} - \rho$ in the AI-dominated case. The constructions imply

$$a \rightarrow \bar{g}, \quad M \rightarrow \left[\frac{\bar{g} \bar{B}^{1-\eta}}{\chi} \right]^{1/\eta} > 0, \quad g_q \rightarrow \bar{g}, \quad g_B \rightarrow 0, \quad M/Y \rightarrow 0.$$

Both transversality expressions in Definition 1 have limiting logarithmic growth

$$-\rho + n + g_K - g_C \rightarrow n - \rho < 0, \quad -r + g_q + g_B \rightarrow \bar{g} - \bar{r} = n - \rho < 0. \quad (\text{A22})$$

Both transversality expressions therefore converge to zero, rather than merely remaining bounded.

Household utility is finite because $\log(C/N)$ grows at most linearly and $\rho > n$, as imposed in the household problem. Log utility is concave, and the household budget is linear at given prices and distributed profits. More explicitly, let $\lambda_t = e^{-\rho t} N_t / C_t$, so $\dot{\lambda}_t = -r_t \lambda_t$ by (5). Concavity of log utility, the household budget, and integration by parts give, for any admissible alternative $(\tilde{C}, \tilde{K})$ with the same initial capital,

$$\int_0^T e^{-\rho t} N_t \log \frac{\tilde{C}_t}{C_t} dt \leq \int_0^T \lambda_t (\tilde{C}_t - C_t) dt = -\lambda_T (\tilde{K}_T - K_T) \leq \lambda_T K_T \rightarrow 0.$$

Here $\tilde{K}_T \geq 0$ and the last limit is the household TVC.

Before applying Lemma 2, I verify that the developer's discounted payoff is well defined. Hold the candidate paths of K, A, N, r fixed, and let $S_t = A_t N_t$ in the positive-share construction and $S_t = K_t$ in the AI-dominated construction. Operating profit increases with efficiency by (20), so $\pi_t(b) \leq \pi_t(\bar{B})$ for every feasible b . Homogeneity implies that $\pi_t(\bar{B})/S_t$ depends only on K_t/S_t and $A_t N_t/S_t$. These ratios converge to $(\bar{k}, 1)$ in the first construction and to $(1, 0)$ in the second. The corresponding limiting static allocations give a finite limit for $\pi_t(\bar{B})/S_t$. Continuity at finite dates therefore gives a constant C_π such that

$$0 \leq \pi_t(b) \leq \pi_t(\bar{B}) \leq C_\pi S_t \quad \text{for every } t \geq 0 \text{ and } b \in [B_0, \bar{B}].$$

Let $D_t = \exp(-\int_0^t r_s ds)$. Since $d \log(D_t S_t)/dt \rightarrow \bar{g} - \bar{r} = n - \rho < 0$, we have $\int_0^\infty D_t S_t dt < \infty$. The bound above therefore makes discounted operating profit uniformly bounded across all feasible alternative research policies. The candidate's discounted research expenditure is also finite, because its research compute is bounded and $r_t \rightarrow \bar{r} > 0$. Its present value of net profit is consequently finite. An alternative with infinite discounted research expenditure has payoff $-\infty$ and cannot improve on it.

It remains to verify condition (A2) in Lemma 2. Shrink the open neighborhood $\mathcal{U}_L$ or $\mathcal{U}_A$ so that

$$\eta \left(1 - \frac{B_0}{\bar{B}} \right) < \min \left\{ \alpha, \frac{1}{\sigma} \right\}.$$

This is possible because $B_0 = \bar{B} - d_0 \tau_0$ in the first construction and $B_0 = \bar{B} - d_0 \tau_0 h_0^{1/\varphi}$ in the second.

For every fixed $0 < \eta < 1$, the left-hand side can therefore be made arbitrarily small while the right-hand side is positive. The strict inequality preserves an open initial-state neighborhood. Together with the finite payoff and the verified TVC, Lemma 2 establishes global developer optimality.

Final-good optimality follows from concavity and constant returns. The pointwise monopoly optimum is global by Lemma 1. All quantities remain finite at every finite date, and market clearing and the original initial conditions hold by reconstruction.

For $\sigma > 1$, at $\bar{B} = \mathcal{B}(\sigma)$ the positive-labor-share limiting point reaches the boundary and a stable eigenvalue in the boundary normalization becomes zero; this proof does not apply. For $\sigma < 1$ and $\bar{B} \leq \mathcal{B}(\sigma)$, it does not rule out equilibria with different long-run ratios.

Step 5. Initial-state sets and the four conclusions.

The initial-state maps are continuous. For each chosen neighborhood, the stocks satisfying $A_0, N_0, K_0 > 0$ and $0 < B_0 < \bar{B}$ whose normalized coordinates lie in that neighborhood therefore form an open set. Each set is nonempty. In the positive-labor-share cases, choose k_0, d_0 close to $\bar{k}, \bar{d}$ and $\tau_0 > 0$ sufficiently small. Then $A_0 N_0 = 1/\tau_0$, $K_0 = k_0/\tau_0$, and $B_0 = \bar{B} - d_0 \tau_0$ are admissible. In the AI-dominated case, choose d_0 near $\bar{d}$ and $h_0, \tau_0 > 0$ sufficiently small. The inverse relations are $A_0 N_0 = 1/\tau_0$, $K_0 = 1/(\tau_0 h_0^{1/\varphi})$, and $B_0 = \bar{B} - d_0 \tau_0 h_0^{1/\varphi}$. Any positive factorization of $A_0 N_0$ supplies A_0, N_0 . These choices also satisfy the strict efficiency restriction in Step 4.

For each initial state in the resulting set, the local graph selects $C_0, q_0 > 0$. These jump values are unique among the trajectories that remain near the indicated normalized limit and converge to it. The argument does not assert global uniqueness.

Steps 1–2 establish the first three cases, and Step 3 establishes the AI-dominated case. Together with the optimality and transversality checks in Step 4, these constructions satisfy Definition 1. $\square$

Proof of Proposition 2. With all parameters other than σ fixed, define

$$H(\sigma) = \mathcal{R}(\bar{B}) + \delta, \quad \kappa = -\frac{1-\alpha}{\alpha} \log \omega_X > 0, \quad D = \rho + \gamma + \delta.$$

Equation (34) gives

$$H'(\sigma) = \frac{\kappa H(\sigma)}{(\sigma-1)^2} > 0,$$

$$H_\infty \equiv \lim_{\sigma \rightarrow \infty} H(\sigma) = \alpha [(1-\alpha)^2 \bar{B} \omega_X]^{(1-\alpha)/\alpha} < \infty.$$

The AI-dominated domain is $H(\sigma) > D$. Since this domain is nonempty and H is strictly increasing, $H_\infty > D$. By equation (40), limiting wage growth is $\bar{g}_w(\sigma) = \gamma + [H(\sigma) - D]/\sigma$, so

$$\bar{g}'_w(\sigma) = \frac{D - H(\sigma) [1 - \kappa\sigma/(\sigma-1)^2]}{\sigma^2}.$$

The factor $1 - \kappa\sigma/(\sigma-1)^2$ increases strictly from $-\infty$ to 1 as σ increases from 1 to infinity. While this factor is nonpositive, $\bar{g}'_w(\sigma) > 0$. Once this factor is positive, $H(\sigma)[1 - \kappa\sigma/(\sigma-1)^2]$ increases strictly

with σ from zero to $H_\infty > D$. The product therefore equals D at exactly one finite $\sigma^* > 1$. At that point the factor is less than one, so $H(\sigma^*) > D$: the critical point lies inside the AI-dominated domain. Limiting wage growth increases with σ below σ^* and decreases above it within that domain. Hence σ^* is its unique maximizer. All other regimes characterized in Proposition 1 have long-run wage growth equal to γ , whereas $\bar{g}_w(\sigma^*) > \gamma$. Hence σ^* uniquely maximizes long-run wage growth across these regimes. $\square$

A.3 Competitive equilibrium and comparison with monopoly

Proof of Proposition 3. Free entry and the common unit cost $1/B$ imply $p_X = 1/B$. Because improvements are immediately available to every supplier and research receives no separate payment, a supplier's net profit is $-M$. Its optimal research expenditure is zero. The RSI technology (17) then gives $B = B_0$, establishing equation (44). The remaining argument constructs and verifies the allocation at this fixed efficiency.

Net production and normalized dynamics. Write $k = K/(AL)$, $c = C/(AL)$, $u = U/(AL)$, and $y = Y/(AL)$. At $p_X = 1/B_0$, final-good firms' demand condition (13) implies $\partial y/\partial u = 1$. Their choice of AI services therefore maximizes output net of compute at a given k . For the proof, denote this net output per unit of effective labor by

$$f(k) = \max_{u \geq 0} \left\{ k^\alpha \left[\omega_L + \omega_X (B_0 u)^{(\sigma-1)/\sigma} \right]^{\sigma(1-\alpha)/(\sigma-1)} - u \right\}. \quad (\text{A23})$$

At each fixed $k > 0$, the marginal product of compute, $\partial y/\partial u$, decreases strictly from infinity to zero as u rises from zero to infinity. The derivative of net output with respect to u is $\partial y/\partial u - 1$, so there is a unique positive maximizer. Its first-order condition is $u = (1 - \alpha)s_X y$. Differentiating this condition with respect to k , using (15), gives the compute elasticity below. The envelope theorem applied to (A23) gives $f'(k)$; differentiating $f'(k)$ then gives the curvature of net production:

$$\frac{d \log u}{d \log k} = \frac{\alpha}{e_X}, \quad f'(k) = \alpha \frac{y}{k}, \quad \frac{k f''(k)}{f'(k)} = -\frac{(1 - \alpha)(1 - s_X)}{\sigma e_X} < 0.$$

Thus f is smooth, increasing, and strictly concave, with $f(0) = 0$ and $f'(0+) = \infty$. As $k \rightarrow \infty$, the optimal u tends to infinity and $s_X \rightarrow 1$. The pricing condition gives $X/Y \rightarrow B_0(1 - \alpha)$, while (32) gives $Z/X \rightarrow \omega_X^{\sigma/(\sigma-1)}$. It follows that

$$\lim_{k \rightarrow \infty} f'(k) = \lim_{k \rightarrow \infty} \frac{f(k)}{k} = \mathcal{R}^{\text{mc}}(B_0) + \delta, \quad f(k) - k f'(k) = (1 - \alpha)(1 - s_X)y > 0. \quad (\text{A24})$$

The limit of $f'(k)$ also yields equation (43).

Set $d = \rho - n > 0$ and $G(k) = f(k) - (\delta + n + \gamma)k$. The household's objective, up to a positive factor and an additive constant, is $\int_0^\infty e^{-dt} \log c_t dt$. The resource constraint (26) with $M = 0$, the

Euler equation (5), and the transversality condition (6) become

$$\dot{k} = G(k) - c, \quad \frac{\dot{c}}{c} = G'(k) - d, \quad \lim_{t \rightarrow \infty} e^{-dt} \frac{k_t}{c_t} = 0. \quad (\text{A25})$$

Subtracting $\dot{k}/k$ from $\dot{c}/c$ in (A25) and substituting the definition of G gives an identity useful for constructing the trajectory:

$$\frac{d}{dt} \log \frac{c}{k} = \frac{c}{k} - d - \frac{f(k) - kf'(k)}{k}. \quad (\text{A26})$$

Efficiency below the threshold. If $\mathcal{R}^{\text{mc}}(B_0) < \rho + \gamma$, there is a unique $\bar{k} > 0$ satisfying $G'(\bar{k}) = d$. The stationary consumption level is $\bar{c} = G(\bar{k}) > d\bar{k} > 0$, by (A24). The Jacobian at this point is

$$\begin{pmatrix} d & -1 \\ \bar{c}f''(\bar{k}) & 0 \end{pmatrix},$$

whose determinant is negative. The stable curve through $(\bar{k}, \bar{c})$ has positive slope, with $\dot{k}, \dot{c} > 0$ on its left branch and $\dot{k}, \dot{c} < 0$ on its right branch.

Local convergence is not enough to establish existence from every initial capital stock. I therefore show that the two branches of this stable curve together cover every $k > 0$. On the left branch, $dk < c < G(k)$; on the right, $c > \max\{dk, G(k)\}$. To verify these restrictions, at $c = dk$ equation (A26) points toward $c/k < d$, from which the path cannot return to $\bar{c}/\bar{k} > d$. At $c = G(k)$, the time derivative of $\dot{k} = G(k) - c$ is $-\dot{c}$. Below $\bar{k}$, crossing this curve makes capital decrease while consumption increases, precluding convergence to $\bar{k}$. Above $\bar{k}$, it makes capital increase while consumption decreases, again precluding convergence. The stable branches cannot make either crossing.

Continuing the left branch backward therefore takes k to zero: otherwise its bounded, monotone coordinates would approach a positive stationary point different from $(\bar{k}, \bar{c})$, which does not exist. The right branch extends to arbitrarily large k . It cannot approach another finite stationary point or diverge in consumption at a finite positive k : along this branch $dc/dk = c(G'(k) - d)/(G(k) - c)$ remains bounded when c becomes large on a compact interval of k . For each $k_0 > 0$, the stable curve consequently supplies a positive trajectory converging to $(\bar{k}, \bar{c})$ and satisfying the transversality condition in (A25).

Efficiency at or above the threshold. If $\mathcal{R}^{\text{mc}}(B_0) \geq \rho + \gamma$, strict concavity implies $G'(k) > d$ and $G(k) > dk$ at every finite $k > 0$. For any $k_0 > 0$, consider initial consumption in $[dk_0, G(k_0)]$ and the region

$$k > 0, \quad dk < c < G(k).$$

I choose initial consumption so that the path stays between these boundaries. At the lower boundary, (A26) points strictly out of the region: consumption falls relative to capital. At the upper boundary, $\dot{k} = 0$ and $\dot{c} > 0$, so consumption rises above $G(k)$. Inside, both k and c increase. There is no

finite-time escape to infinity because concavity bounds $G(k)$ above by an affine function.

The endpoint choices of initial consumption leave through opposite boundaries. A small change in initial consumption preserves a finite-time exit through either boundary, by continuous dependence of the solution on its initial value. Thus the choices that first exit through the lower boundary and those that first exit through the upper boundary form two nonempty, disjoint sets that are open within the interval. They cannot cover the whole interval $[dk_0, G(k_0)]$, because an interval is connected. At least one initial consumption choice therefore generates a path that never leaves the region.

Along this trajectory $k \rightarrow \infty$. A finite limit would also bound c and require a positive stationary point, but $G'(k) > d$ excludes one. Moreover, $d < c/k < G(k)/k$, and the upper bound satisfies

$$\frac{G(k)}{k} \rightarrow \mathcal{R}^{\text{mc}}(B_0) - n - \gamma.$$

If $\mathcal{R}^{\text{mc}}(B_0) = \rho + \gamma$, these bounds imply $c/k \rightarrow d$. If $\mathcal{R}^{\text{mc}}(B_0) > \rho + \gamma$, the same limit follows from (A26): $(f - kf')/k \rightarrow 0$, and any persistent excess of c/k over d would make c/k increase until it exceeded its upper bound $G(k)/k$. More precisely, for any $\varepsilon > 0$ and sufficiently large t , reaching $c/k \geq d + \varepsilon$ implies $d \log(c/k)/dt \geq \varepsilon/2$ thereafter until that contradiction occurs. Thus

$$\frac{c}{k} \rightarrow d, \quad \frac{\dot{k}}{k} \rightarrow \mathcal{R}^{\text{mc}}(B_0) - \rho - \gamma. \quad (\text{A27})$$

The transversality condition holds because $c/k > d$. At $\mathcal{R}^{\text{mc}}(B_0) = \rho + \gamma$, capital per effective worker diverges but its growth rate tends to zero.

Optimality, uniqueness, and decentralization. The preceding constructions satisfy global optimality, not only the Euler equation. Let $\lambda_t = e^{-dt}/c_t$ be discounted marginal utility along a constructed path; then $\dot{\lambda} = -G'(k)\lambda$. For any alternative feasible path $(\tilde{k}, \tilde{c})$ with the same k_0 , concavity of log utility and G , the resource constraint in (A25), and integration by parts give

$$\int_0^T e^{-dt} (\log \tilde{c} - \log c) dt \leq -\lambda_T (\tilde{k}_T - k_T) \leq \lambda_T k_T \rightarrow 0.$$

The constructed path has finite utility: below the efficiency threshold, c converges to a positive constant; at or above the threshold, c is positive, increasing, and at most exponentially growing. Thus the inequality verifies optimality in the fixed- B_0 allocation problem. Strict concavity of log utility and concavity of the resource constraint imply uniqueness of consumption; the state equation then gives uniqueness of capital. This verification does not allow a planner to improve B : it concerns the allocation with the technology fixed by the private research choices.

The maximizing u in (A23) determines $U = (AL)u$. This compute use, the factor prices in (11) and (12), and $p_X = 1/B_0$ decentralize this allocation. In particular, $wN = (AL)[f(k) - kf'(k)]$, so the household budget coincides with the resource constraint. The household's Euler and transversality conditions verify household optimality at these prices. Strict concavity also implies that any other competitive equilibrium would solve the same fixed-technology allocation problem and hence have

the same aggregate allocation.

Growth rates and income shares. Below the threshold, convergence to $(\bar{k}, \bar{c})$ gives $r \rightarrow \rho + \gamma$, $g_{Y/N}, g_w \rightarrow \gamma$, and a positive limiting labor share, because optimal u and s_X have finite interior limits. At or above the threshold, $k, u \rightarrow \infty$ and $s_X \rightarrow 1$. Differentiating the production and wage equations (7)–(8) and (12), using $X = B_0 U$ and $d \log u / d \log k = \alpha / e_X$, gives

$$\frac{d \log y}{d \log k} = 1 - \frac{(1 - \alpha)(1 - s_X)}{\sigma e_X} \rightarrow 1, \quad \frac{d \log(w/A)}{d \log k} = \frac{\alpha}{\sigma e_X} \rightarrow \frac{1}{\sigma}.$$

Combining these elasticities with (A27) proves the growth limits in part (ii). Equation (A24) gives the interest-rate limit, and $wL/Y = (1 - \alpha)(1 - s_X) \rightarrow 0$ gives the labor-share limit.

At the boundary $B_0 = \mathcal{B}^{\text{mc}}(\sigma)$, the same calculations give $g_{Y/N}, g_w \rightarrow \gamma$, $r \rightarrow \rho + \gamma$, and $wL/Y \rightarrow 0$. The labor share therefore vanishes without a permanent growth premium. $\square$

Proof of Proposition 4. Fix $\sigma > 1$. The limiting returns in equations (34) and (43), derived above, satisfy

$$\mathcal{R}^{\text{mc}}(B) = \mathcal{R}(B/(1 - \alpha)).$$

Both returns increase strictly with B . Since each threshold is the efficiency at which the corresponding return equals $\rho + \gamma$,

$$\mathcal{R}^{\text{mc}}((1 - \alpha)\mathcal{B}(\sigma)) = \mathcal{R}(\mathcal{B}(\sigma)) = \rho + \gamma.$$

Uniqueness of the competitive efficiency threshold gives $\mathcal{B}^{\text{mc}}(\sigma) = (1 - \alpha)\mathcal{B}(\sigma) < \mathcal{B}(\sigma)$, establishing equation (46). $\square$

Proof of Corollary 1. Proposition 4 ensures that the interval in equation (47) is nonempty. Fix $\bar{B}$ in this interval. The labor-bottleneck construction of Proposition 1(iii) uses initial values satisfying equation (A17), with $B_0 = \bar{B} - d_0 \tau_0$, d_0 near $\bar{d} > 0$, and $\tau_0 = 1/(A_0 N_0) > 0$ arbitrarily small. Since $\bar{B} > \mathcal{B}^{\text{mc}}(\sigma)$, I can restrict this construction to $B_0 > \mathcal{B}^{\text{mc}}(\sigma)$ while preserving a nonempty open set of initial states.

From each state in this set, the constructed monopoly equilibrium has $B \rightarrow \bar{B}$ and $g_{Y/N}, g_w \rightarrow \gamma$. Proposition 3 supplies a competitive equilibrium from the same state. Since $B_0 > \mathcal{B}^{\text{mc}}(\sigma)$, this equilibrium is AI-dominated, with $B = B_0$ and both limiting growth rates strictly above γ . This proves the comparison without requiring monopoly equilibrium existence from arbitrary initial states or global uniqueness under monopoly.

The accompanying comparison of AI services follows from the CES share identity in equation (A8) and $L = N$. Under monopoly, s_X converges to an interior value, so $X/(AL)$ converges to a positive constant. Under competition, $s_X \rightarrow 1$; since $\sigma > 1$, the same identity gives $X/(AL) \rightarrow \infty$. The common paths of A and L therefore imply that AI services eventually exceed their monopoly level under competition, without establishing their ordering at every date. $\square$

Proof of Corollary 2. Equations (34) and (43) imply

$$\mathcal{R}(\bar{B}) = \mathcal{R}^{\text{mc}}((1 - \alpha)\bar{B}).$$

Since $\mathcal{R}^{\text{mc}}$ is strictly increasing in AI efficiency, it follows that

$$\mathcal{R}(\bar{B}) \geq \mathcal{R}^{\text{mc}}(B_0) \iff (1 - \alpha)\bar{B} \geq B_0,$$

with equality on one side if and only if equality holds on the other. Under the two AI-dominated equilibria, these returns are the limiting interest rates. Equations (39), (40), and (45) express limiting output-per-worker and wage growth as strictly increasing functions of the limiting interest rate, with slopes one and $1/\sigma$, respectively. Since the parameters are common to both economies, the same ordering applies to both growth rates, including the equality case. $\square$

B Numerical algorithm and equilibrium audit

B.1 Numerical solution

For the RSI-activation comparisons in Sections 6.2 and 6.3, I compute numerical approximations to equilibrium trajectories by solving a boundary-value problem (BVP). Initial capital K_0 and AI efficiency B_0 are given, whereas initial consumption C_0 and the developer's shadow price q_0 are determined jointly with the transition. I solve the equilibrium differential equations (28) over a finite horizon, using the given stocks as initial conditions and the long-run characterization in Proposition 1 to construct terminal conditions.

For each σ , I obtain the pre-RSI capital stock from Equation (49) under monopoly pricing with efficiency fixed at B_0 . For the competitive-to-monopoly comparison in Section 6.4, I instead impose $r = \rho + \gamma$ with marginal-cost pricing $p_X = 1/B_0$, giving $K_0 = 6.535$ at $\sigma = 1.5$. In both cases, the resource constraint with $M = 0$ and $\dot{K} = (n + \gamma)K$ gives pre-event consumption. The post-event monopoly BVP inherits the stocks but selects consumption, research, and the shadow price anew.

In the normalized coordinates of Appendix A.2, I require the terminal deviation from the limiting point to lie in the stable subspace of the linearized dynamics. The resulting two terminal restrictions, together with the initial stocks, select initial consumption and the shadow price without requiring the economy to reach its long-run limit at a finite date.

I solve this BVP with `scipy.integrate.solve_bvp` from the open-source Python library SciPy (Virtanen et al., 2020). The solver implements fourth-order collocation with residual control similar to that of Kierzenka and Shampine (2001). At each trial state, I solve the static monopoly problem and recover research expenditure from its first-order condition.

B.2 Accuracy and equilibrium checks

I check the dynamic and first-order conditions independently of the solver’s residual report, tighten tolerances, and extend each horizon twice. I also check pre-event conditions, continuity at RSI activation, proximity to the limiting equilibrium, and the asymptotic decay of both transversality expressions at rate $n - \rho < 0$. The [digital edition](#) reports the computational coordinates, continuation procedure, tolerances, and accuracy table. All plotted values lie within the solved interval. The competitive-to-monopoly paths satisfy the same checks.

Small equation residuals do not establish global developer optimality. All eight RSI-activation configurations have $\alpha = 0.33$, $\eta = 0.20$, and $B_0/\bar{B} = 0.01$. For each of the four elasticities $\sigma \in \{0.90, 1.00, 1.10, 1.50\}$, condition (A2) holds because

$$\eta \left(1 - \frac{B_0}{\bar{B}} \right) = 0.20(1 - 0.01) = 0.198 < 0.33 = \min \left\{ \alpha, \frac{1}{\sigma} \right\}.$$

This condition does not depend on the dated values of capital and effective labor or on research productivity χ . The baseline competitive-to-monopoly comparisons use the same parameters and initial efficiency ratio, so the same calculation applies. The supplementary growth-reversal exercise instead has $B_0/\bar{B} = 0.90$, giving $0.20(1 - 0.90) = 0.02 < 0.33$.

The long-run limits used to construct the terminal conditions imply a finite present value of net profit for any feasible trajectory attaining those limits. Let $D_t = \exp(-\int_0^t r_s ds)$. Discounted output $D_t Y_t$ has limiting logarithmic growth $n - \rho < 0$, and operating profit is bounded above by output. Research expenditure converges to a finite level, while the interest rate converges to a positive limit. Both discounted operating profit and discounted research expenditure are therefore integrable.

Together with the parameter restriction verified above, these properties allow Lemma 2 to establish global developer optimality for a feasible trajectory satisfying the research condition, costate equation, and transversality condition. The numerical checks assess how closely the computed paths satisfy these conditions over the solved horizon and approach the required long-run limits. They do not prove equilibrium existence from arbitrary initial stocks.

Replication files. I use NumPy for array operations and linear algebra ([Harris et al., 2020](#)) and Matplotlib to generate the figures ([Hunter, 2007](#)). The code and simulation data are available in the [repository](#). The [replication guide](#) records package versions, implementation details, numerical tolerances, and commands for checking and reproducing the results, with links to the additional institutional experiments.

C Equilibrium without an upper bound on AI efficiency

I now return to the monopoly economy, remove the upper bound on AI efficiency, and compare the three substitution regimes. The RSI technology becomes

$$\dot{B} = \chi(BM)^\eta. \quad (\text{C1})$$

All other primitives remain as in Section 3. In Definition 1, I set $\psi(B) = 1$ and $\psi'(B) = 0$ and remove the upper restriction on B . Equilibrium still requires finite allocations at every finite date, global agent optimality, and both transversality conditions.

The results differ in scope across substitution regimes. Under complementarity, long-run average output-per-worker growth cannot exceed the rate of labor-augmenting technological progress, even without an upper bound on AI efficiency. At unit elasticity, I give sufficient parameter restrictions for an analytically tractable balanced-growth equilibrium. Under substitution and sufficiently strong returns to research, the developer can make discounted net profit arbitrarily large even over a fixed finite horizon, ruling out an equilibrium.

C.1 Complementarity

When $\sigma < 1$, effective labor remains necessary even if AI becomes arbitrarily efficient. Equation (8) gives an upper bound for Z and Y which does not depend on B :

$$Z \leq \omega_L^{\sigma/(\sigma-1)} AL, \quad (\text{C2})$$

$$Y \leq \omega_L^{\sigma(1-\alpha)/(\sigma-1)} K^\alpha (AL)^{1-\alpha}. \quad (\text{C3})$$

Proposition 5 (Production bounds under complementarity). *Let $\sigma < 1$ and $\bar{B} = \infty$. Along any feasible path satisfying the production equations (7) and (8), labor-market clearing $L = N$ in equation (27a), and the resource constraint (26), long-run average output-per-worker growth cannot exceed γ :*

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log \left(\frac{Y_t/N_t}{Y_0/N_0} \right) \leq \gamma. \quad (\text{C4})$$

Proof. See the [proof in the digital edition](#). □

Complementarity bounds long-run average output-per-worker growth despite RSI and the absence of an AI-efficiency upper bound. This does not prevent growth from temporarily exceeding γ .

C.2 Unit elasticity

At unit elasticity, production becomes

$$Y = K^\alpha (AL)^{(1-\alpha)\omega_L} X^{(1-\alpha)\omega_X}. \quad (\text{C5})$$

Consider a balanced-growth path (BGP) with positive constant K/Y , C/Y , and M/Y . The AI-demand and monopoly conditions, equations (13) and (19), imply $U/Y = (1 - \alpha)^2 \omega_X^2$. Along the BGP, $g_K = g_U = g_M = g_Y$. Substituting $X = BU$ into production, equation (C5), and log-differentiating this equation and the RSI technology in equation (C1) yields

$$g_Y = n + \gamma + \frac{\omega_X}{\omega_L} g_B, \quad (\text{C6})$$

$$(1 - \eta)g_B = \eta g_Y. \quad (\text{C7})$$

The first relation describes how AI improvement raises production growth. The second describes how expanding research resources sustain AI improvement. Solving these relations gives candidate growth rates. The next proposition establishes that they belong to an equilibrium.

Proposition 6 (An uncapped balanced-growth equilibrium and local convergence). *Suppose $\sigma = 1$, $\bar{B} = \infty$, and*

$$\eta + \omega_X < 1. \quad (\text{C8})$$

(i) *There is an equilibrium trajectory such that*

$$g_{Y/N} = g_w = \gamma + \frac{\eta \omega_X}{1 - \eta - \omega_X} (n + \gamma), \quad (\text{C9})$$

$$r = \rho + \gamma + \frac{\eta \omega_X}{1 - \eta - \omega_X} (n + \gamma), \quad (\text{C10})$$

$$\frac{wL}{Y} = (1 - \alpha)\omega_L > 0, \quad (\text{C11})$$

$$g_B = \frac{\eta(1 - \omega_X)}{1 - \eta - \omega_X} (n + \gamma). \quad (\text{C12})$$

(ii) *There is an open neighborhood of the initial condition (A_0, N_0, K_0, B_0) of the BGP in part (i) such that, from every initial condition in that neighborhood, there exists an equilibrium trajectory converging to the corresponding BGP in normalized variables.*

Proof. See the [proof in the digital edition](#). □

Proposition 6 shows that RSI can sustain permanently faster wage growth without a vanishing labor income share. Equations (C9) and (C11) imply $g_w = g_{Y/N} > \gamma$ while $wL/Y = (1 - \alpha)\omega_L$ remains constant and positive. Wages therefore keep pace with output per worker. This contrasts with the regimes with an upper bound in Proposition 1, where a permanent wage-growth premium occurs only in the AI-dominated regime and accompanies a vanishing labor share. A declining labor share is therefore not a necessary consequence of recursive self-improvement.

Growth in effective labor expands the resources available for research, which raises B and further expands output. The common denominator $1 - \eta - \omega_X$ captures this reinforcement: a higher η strengthens the research feedback, while a higher ω_X makes output growth more responsive to AI improvement. A related dependence appears in Jones (1995), where long-run growth depends on an expanding research workforce.

C.3 Substitution

With $\sigma > 1$, Proposition 1 establishes AI-dominated equilibria whose limiting interest rate is $\mathcal{R}(\overline{B})$, which grows without bound as $\overline{B} \rightarrow \infty$. The same function also provides a lower bound on the equilibrium interest rate at each date without an efficiency upper bound:

$$r_t \geq \mathcal{R}(B_t). \quad (\text{C13})$$

To see this, combine equations (31), (32), and (33) before taking the limit $s_X \rightarrow 1$:

$$\frac{Y}{K} = \left[B(1 - \alpha)\omega_X^{\sigma/(\sigma-1)}(1 - e_X)s_X^{-1/(\sigma-1)} \right]^{(1-\alpha)/\alpha}. \quad (\text{C14})$$

For fixed B and $\sigma > 1$, this expression decreases with s_X and approaches its lower limit as $s_X \rightarrow 1$. Substitution into $r = \alpha Y/K - \delta$ gives inequality (C13). Thus $\mathcal{R}(B)$ is the lowest return compatible with the production and monopoly conditions at that efficiency level, attained in the zero-labor-share limit. If B grows without bound along a trajectory, inequality (C13) requires $r \rightarrow \infty$ and $Y/K \rightarrow \infty$. Production also gives $Z/Y = (Y/K)^{\alpha/(1-\alpha)} \rightarrow \infty$, while the Euler equation (5) requires $g_{C/N} = r - \rho \rightarrow \infty$.

Davidson et al. (2026) show that technological and economic feedback loops can overcome diminishing returns to research and generate *hyperbolic growth*, with variables diverging at a finite date. This *finite-time singularity* is stronger than an indefinitely increasing growth rate. Their benchmark fixes saving and factor-allocation shares; I therefore cannot apply its explosion conditions directly to the optimal choices studied here.

Under Definition 1, a finite-time singularity cannot constitute an infinite-horizon equilibrium, which requires finite allocations at every finite date. The following proposition shows that, when research returns are sufficiently strong, the developer can make discounted net profit arbitrarily large even over a fixed finite horizon.

Proposition 7 (Unbounded discounted net profit). *Suppose $\sigma > 1$, $\overline{B} = \infty$, and $\eta > \alpha$. Fix an initial efficiency B_0 , a finite horizon $T > 0$, continuous paths of K, A, L , and an integrable interest-rate path. The developer's discounted net profit in problem (21), restricted to $[0, T]$, is unbounded above over feasible research trajectories. Consequently, no equilibrium satisfying Definition 1 exists under these parameter restrictions.*

Proof. See the [proof in the digital edition](#). □

Declaration of AI use

During the preparation of this work, I used OpenAI Codex to assist with mathematical derivations, consistency checks, manuscript editing, and the preparation and testing of computational code. I also used Anthropic’s Claude to audit the manuscript. I take full responsibility for the content of the paper.

References

- Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.
- 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 237–282. University of Chicago Press.
- Autor, D. H. and Kausik, B. N. (2026). Resolving the automation paradox: Falling labor share, rising wages. arXiv:2601.06343.
- Beaudry, P. and Portier, F. (2014). News-driven business cycles: Insights and challenges. *Journal of Economic Literature*, 52(4):993–1074.
- Berg, A., Buffie, E. F., and Zanna, L.-F. (2018). Should we fear the robot revolution? (The correct answer is yes). *Journal of Monetary Economics*, 97:117–148.
- 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.
- Chow, T., Halperin, B., and Mazlish, J. Z. (2026). Transformative AI, existential risk, and real interest rates. Working paper; first posted January 2023.
- 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, National Bureau of Economic Research.
- Demirer, M., Fradkin, A., Tadelis, N., and Peng, S. (2025). The emerging market for intelligence: Pricing, supply, and demand for LLMs. NBER Working Paper 34608, National Bureau of Economic Research.
- Duffy, J. and Papageorgiou, C. (2000). A cross-country empirical investigation of the aggregate production function specification. *Journal of Economic Growth*, 5(1):87–120.
- Dyatlov, S. (2018). Notes on hyperbolic dynamics.

- Erdil, E. and Besiroglu, T. (2022). Algorithmic progress in computer vision. *arXiv preprint arXiv:2212.05153*.
- 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:2503.04941*.
- Feldman, W. M. (2021). Stable manifold theorem. MATH 6410: Ordinary Differential Equations, University of Utah. Fall lecture slides.
- Gans, J. S. (2024). Market power in artificial intelligence. NBER Working Paper 32270, National Bureau of Economic Research.
- Gans, J. S. (2026). Growth in AI knowledge. Working Paper 5282272, SSRN. Revised August 5, 2026.
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.
- Hémous, D. and Olsen, M. (2022). The rise of the machines: Automation, horizontal innovation, and income inequality. *American Economic Journal: Macroeconomics*, 14(1):179–223.
- Hernandez, D. and Brown, T. B. (2020). Measuring the algorithmic efficiency of neural networks. *arXiv preprint arXiv:2005.04305*.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3):90–95.
- Jones, C. I. (1995). R&D-based models of economic growth. *Journal of Political Economy*, 103(4):759–784.
- Jones, C. I. (2026). AI and our economic future. *Journal of Economic Perspectives*, 40(3):3–22.
- 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, May 2026.
- Jones, L. E. and Manuelli, R. E. (1990). A convex model of equilibrium growth: Theory and policy implications. *Journal of Political Economy*, 98(5, Part 1):1008–1038.
- Kierzenka, J. and Shampine, L. F. (2001). A BVP solver based on residual control and the MATLAB PSE. *ACM Transactions on Mathematical Software*, 27(3):299–316.
- Korinek, A. and McKelvey, P. (2026). Measuring the AI economy. Staff Working Paper 2026-20, Bank of Canada.
- Korinek, A. and Ng, D. X. (2019). Digitization and the macro-economics of superstars. Working paper, December 2019 version.

- Korinek, A. and Vipra, J. (2025). Concentrating intelligence: Scaling and market structure in artificial intelligence. *Economic Policy*, 40(121):225–256.
- Kremer, M. (1998). Patent buyouts: A mechanism for encouraging innovation. *The Quarterly Journal of Economics*, 113(4):1137–1167.
- Kydland, F. E. and Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–491.
- Liu, T. and Wan, R. (2026). AGI, ANI and economic growth. Working paper, SSRN 6776339.
- OECD (2025). OECD Global Long-Run Economic Scenarios: 2025 Update. Technical Report 36, OECD Publishing, Paris.
- Palivos, T. and Karagiannis, G. (2010). The elasticity of substitution as an engine of growth. *Macroeconomic Dynamics*, 14(5):617–628.
- Ray, D. and Mookherjee, D. (2022). Growth, automation, and the long-run share of labor. *Review of Economic Dynamics*, 46:1–26.
- Rebelo, S. (1991). Long-run policy analysis and long-run growth. *Journal of Political Economy*, 99(3):500–521.
- Restrepo, P. (2025). We won’t be missed: Work and growth in the AGI world. NBER Working Paper 34423, National Bureau of Economic Research.
- Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2):S71–S102.
- Sastry, G., Heim, L., Belfield, H., et al. (2024). Computing power and the governance of artificial intelligence. *arXiv preprint arXiv:2402.08797*.
- Trammell, P. and Korinek, A. (2023). Economic growth under transformative AI. NBER Working Paper 31815, National Bureau of Economic Research. Revised April 2026.
- United Nations (2024). World Population Prospects 2024: Summary of Results. Technical Report UN DESA/POP/2024/TR/NO. 9, United Nations, Department of Economic and Social Affairs, Population Division, New York.
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272.