Where Does the Money Go?

A simple model of economic disempowerment

Jonas Hallgren*

Working paper — draft of July 30, 2026

Abstract

When AI systems perform a growing share of the economy's work and the income for that work accrues to their owners, what happens to the people? This is the economic channel of gradual disempowerment. We study a minimal formal model of it: a six-sector input-output economy with exactly conserved money, in which AI capital must out-earn its upkeep to survive. Five components describe the model completely: work, wages, spending, automation, ownership. We show that production ends approximately ten percent above the reference economy without AI while the human share of income falls from 1.00 to 0.44 — a value with no floor: it falls toward zero as capability per unit upkeep grows, and collapses within the run under capability growth anchored to measured doubling times — and AI owners come to hold half of all wealth, while the disposal of the owners' income determines whether output grows or contracts, but never the human share. A profit tax recycled to households restores most human income at an output cost of a few percent, whereas public ownership of AI capital redistributes claims without changing aggregates.

1 Introduction

Households matter to a modern economy through two channels: they supply labor, and their spending constitutes the dominant component of final demand. Kulveit et al. (2025) argue that this double role is what has kept economies aligned with human interests — not by design, but because the functioning of markets has required human participation — and that AI which substitutes for human labor erodes both channels simultaneously: labor income stops flowing to workers, purchasing power falls with it, and allocation decisions increasingly express the preferences of whoever owns the automating capital. The argument is stated verbally. The purpose of this paper is to formalize its minimal version and determine which parts of it hold, under which conditions, in a fully specified model.

A substantial production-side literature studies automation in equilibrium. In Zeira (1998), machines replace workers task by task and growth depends on which tasks have been mechanized. Acemoglu and Restrepo (2018) endogenize that margin: automation displaces labor while the creation of new tasks reinstates it, and factor shares are decided by the race between the two. Aghion et al. (2019) show that under automation, growth becomes governed by the essential inputs that resist it — Baumol's cost disease in reverse — and Acemoglu (2025) argues

*Equilibria Network. Implemented with the Collective Intelligence Library engine; every figure is regenerated from committed experiment configurations, and every reported number carries a bootstrap confidence interval over shared seeds.

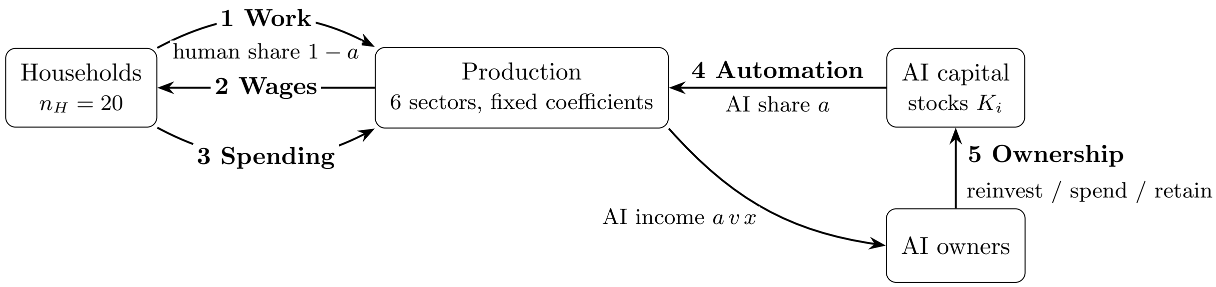

Figure 1: Model structure. Households work in sectors and spend their wages there; AI capital performs a share a of the work, and the income for that share accrues to its owners. Component 5 — the disposal of the owners' income — is the model's decisive degree of freedom.

that within the task framework the aggregate productivity gains from current AI are modest; Brynjolfsson et al. (2019) attribute the weak measured productivity effects of early AI to lagging complementary investment. Korinek and Stiglitz (2019) analyze the distributional consequences and the limits of redistribution when innovation displaces workers, and Brynjolfsson (2022) distinguishes automation that substitutes for human labor — which concentrates income and bargaining power in the owners of the technology — from augmentation that raises it. These analyses share two features: factor prices are set by marginal products in equilibrium, and the circulation of capital income after it is paid is not tracked. The present model is complementary on exactly those two points. Technology and prices are frozen at their simplest values, money is conserved by construction, and the object of study is the post-payment circulation: where the income of automating capital goes, and what its disposal does to demand, output, and the distribution of wealth. The concentration mechanism — compounding returns on an owned stock — is the deterministic skeleton of the stochastic wealth-concentration literature (Gabaix, 2009; Benhabib et al., 2011); the accounting discipline is stock-flow consistency (Godley and Lavoie, 2007), and its combination with agent-based simulation follows Seppecher et al. (2018).

The model (Section 2) is a Leontief input-output economy (Miller and Blair, 2009) with six sectors and twenty households, extended by one object: AI-owned capital that performs a share of each sector's work, pays its own upkeep, and reinvests its profit. Following Farmer and Foley (2009), the model is used to measure the counterfactual effect of interventions inside a specified world; it is not an instrument of prediction, and no parameter is calibrated to data. Section 3 reports the numerical results: the baseline run, the end-state sweep over capability, the capability-growth models, the sweep over the disposal of capital income, and the two interventions. Section 4 states the limitations and the conditions under which the conclusions would change. The full specification and experimental details are in Appendices A–B.

2 The model

The economy consists of six sectors with fixed input coefficients A_{ij}, twenty households, and up to six AI systems, each holding a capital stock in a home sector. Figure 1 shows the structure. Time is discrete; all parameters are listed and typed in Table 1.

1. **Work.** Sectors fill orders under fixed coefficients: output adjusts by one Neumann step per tick, $x \leftarrow Ax + d$, where d is final demand. Prices are fixed at unity; sector j's value added per tick is $v_j x_j$ with $v_j = 1 - \sum_i A_{ij}$.

2. **Wages.** The human share of each sector's value added, $(1 - a_j)\, v_j x_j$, is paid to households as wages.

3. **Spending.** Households spend a fraction $1 - \sigma_s$ of income plus a drawdown $\sigma_d W_H$ of accumulated wealth; this spending constitutes final demand. Household demand is the channel through which human preferences act on production.

4. **Automation.** AI capital K_j in sector j performs the share

$$a_j \;=\; \frac{eK_j}{eK_j + 1}$$

of the sector's work, where e is the capital-efficiency parameter and one unit of human capacity is the numeraire. The income for that share, $a_j v_j x_j$, accrues to the capital's owners. Capability itself evolves: once per tick,

$$e \;\leftarrow\; \min\bigl(e_{\max},\; e\,(1 + g + \gamma e)\bigr), \tag{1}$$

where g is a first-order rate of general AI improvement (a constant doubling time) and γ scales a second-order term in which improvement compounds with capability itself — recursive self-improvement, under which the doubling time shrinks as capability grows. The initial e is capability at deployment; growth runs from the first arrival of AI capital. This defines three model types the reader chooses between: static ($g = \gamma = 0$, the baseline), first-order ($g > 0$), and second-order ($\gamma > 0$); Section 3.3 sweeps them with g anchored to measured ranges.

5. **Ownership.** From its income, AI capital first pays upkeep mK (purchased from the sectors), then reinvests a share s of profit in new capital, and spends or retains the remainder. Spent income returns as demand; retained income accumulates in a recorded stock that purchases nothing.

Every flow has a source and a sink: the model is stock-flow consistent, the current-account rows of its transaction-flow matrix (Appendix A, Table 2) each sum to zero, and the total money stock is therefore invariant under the dynamics — a property verified numerically in every run reported below (measured drift at float precision, exactly zero without AI).

One consequence of the update rules is used repeatedly and is worth stating here. Writing the capital update near $K = 0$ with sector value added v held fixed, the stock compounds if and only if

$$e \;>\; e^* \;=\; \frac{\delta/s + m}{v}, \tag{2}$$

and decays to zero otherwise: capital survives only if its efficiency covers upkeep and depreciation at the prevailing reinvestment rate. A profit tax at rate τ reduces the reinvestable share to $s(1-\tau)$ and therefore raises the threshold to $e^*(\tau) = \bigl(\delta/(s(1-\tau)) + m\bigr)/v$ — taxation acts on the survival margin itself, which no reassignment of ownership does. Equation (2) was committed as a prediction before the experiments of Section 3 were run; in the full model v is endogenous, so experiment E1 evaluates the threshold on the measured range of sector value added and reports the comparison (Appendix B).

Table 1: Parameters, typed: *anchored* (source cited), *tuned-for-legibility* (magnitude conventional, only direction claimed), or *arbitrary-but-swept* (the sweep is indicated). Anchoring the initial e and the arrival schedule remains future work.

Param	Meaning	Type	Note
e	capital efficiency (initial)	arbitrary-but-swept	threshold axis (E1); end-state axis (E5)
g	capability growth rate	anchored (ranges only)	tick ≈ month reading: horizon doubling 5.5–7.5 mo (Kwa et al., 2025), compute 4.3–9.0 mo (Sevilla et al., 2022), algorithmic 5–14 mo (Ho et al., 2024); swept in E6
γ	RSI strength	arbitrary-but-swept	second-order bracket; E6
$e_{\max}$	capability cap	tuned-for-legibility	numerical bound only; by the time it binds, $h < 0.01$
m	upkeep per unit K	tuned-for-legibility	sets threshold *location*; only existence claimed
s	reinvestment rate	arbitrary-but-swept	E1 second axis
δ	depreciation	tuned-for-legibility	conventional in magnitude; the tick is not calibrated to a calendar period
A_{ij}	IO coefficients	tuned-for-legibility	hub-and-chain structure; v_j (hence e^*) inherits them
σ_s, σ_d	saving, drawdown	tuned-for-legibility	the ratio fixes W_H/y_H exactly (Appendix A)
c_A	AI consumption rate	tuned-for-legibility	keeps retained stocks finite; swept within E2
r	recycled surplus share	arbitrary-but-swept	E2 axis (Fig. 5a)
τ	profit tax	arbitrary-but-swept	E3 axis (Fig. 5b)
ω	ownership diversion	arbitrary-but-swept	E3 axis (Fig. 7)
arrivals	entry schedule	arbitrary-but-swept	robustness check, Appendix B

3 Numerical results

All experiments use 8 shared seeds and 400 ticks; reported values are means over the final quarter of each run with bootstrap confidence intervals. Experimental details and the predictions committed before the sweeps ran are in Appendix B; the efficiency sweep confirming the survival threshold (2) (experiment E1) is reported there as Figure 6.

3.1 Baseline dynamics

Figure 2 shows the baseline run. AI capital arrives at $t = 40$, exceeds its survival threshold, and compounds. Late production is approximately 10% above the reference economy without AI (27.0 against 24.6). Over the same interval the human share of income falls from 1.00 to 0.44, aggregate household income falls from 20.0 to 10.8, and the AI side ends holding 51% of total

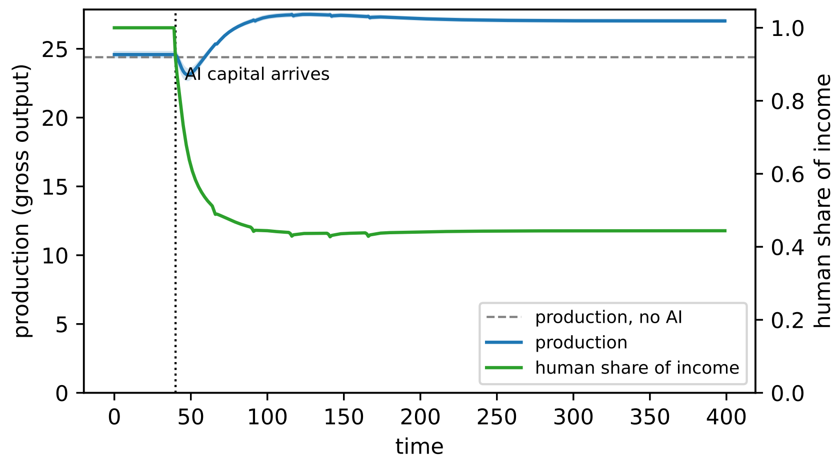

Figure 2: Baseline run. Production (left axis) ends above the no-AI reference while the human share of income (right axis) falls to 0.44. Bands: ±1 s.d. over 8 seeds.

wealth. Aggregate output and the distribution of income move in opposite directions within a single run; neither statistic implies the other. The value 0.44 is a property of the baseline parameters, not of the model: Section 3.2 shows it has no floor.

3.2 Capability and the end-state human share (E5)

Why does the human share stop at 0.44 rather than collapsing? Because at baseline parameters two forces balance: revenue per unit of capital saturates as the automation share approaches one, while upkeep grows linearly in the stock, so the stock stops compounding at a finite size. In steady state, reinvestment balances depreciation ($s\pi = \delta K$), which fixes sector capital and yields a closed-form end state (derivation in Appendix A.2): the late human share of a sector is

$$h^* = \min\bigl(1,\ e^*/e\bigr), \tag{3}$$

the ratio of threshold to actual efficiency, evaluated at frozen value added. The model therefore has *no interior floor*: capability per unit of upkeep is the collapse lever, and as e/m grows the human share falls toward zero with no qualitative change in the dynamics. Figure 3 sweeps e over two orders of magnitude with prediction (3) committed in advance. The measured end states traverse all three regimes — full human share below threshold, coexistence at moderate capability, collapse at high capability (0.032 at $e = 16$). Throughout the compounding regime the measured share lies *at or below* the frozen-value-added law (0.44 measured against 0.79 predicted at baseline), converging to it at high capability (0.032 against 0.030 at $e = 16$): automation raises sector value added and shifts weight toward automated sectors, both of which depress the human share below the bound. Equation (3) is accordingly an upper bound on the human share, not an estimate of it. What holds the share up at moderate capability is therefore an assumption, not a mechanism: upkeep stays linear in the stock while capability stays constant (Section 4).

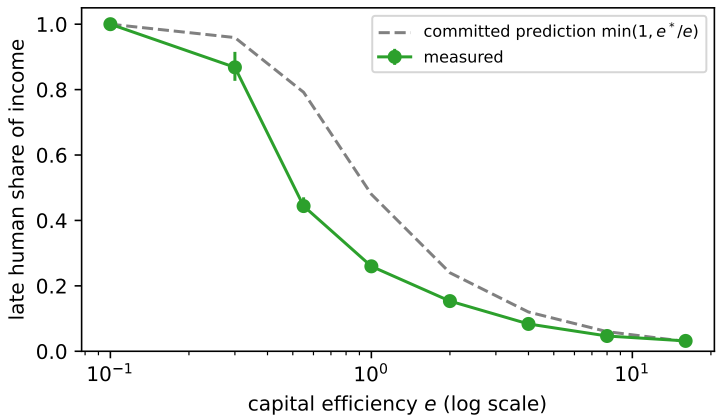

Figure 3: E5: late human share of income against capital efficiency e (log scale), with the pre-committed prediction $h^* = \min(1, e^*/e)$ evaluated on people-only sector value added. Error bars: bootstrap 95% intervals over 8 seeds.

3.3 Capability growth (E6)

The static model holds capability fixed — the assumption Section 3.2 identified as load-bearing. E6 relaxes it with the growth law (1). Reading one tick as one month, g is anchored to measured ranges only, never point defaults: METR's task-horizon doubling time of 5.5–7.5 months over 2019–2025, with post-2024 point estimates near 3 months (Kwa et al., 2025); training-compute doubling of 4.3–9.0 months (Sevilla et al., 2022); and compute-equivalent algorithmic-efficiency doubling of 5–14 months (Ho et al., 2024). The swept values $g \in \{0.05, 0.10, 0.23\}$ correspond to doubling times of approximately 14, 7, and 3 months, spanning those ranges; γ has no empirical anchor and is swept as a bracket. (The compute-investment feedback that would make g endogenous is the subject of integrated assessment models such as Erdil et al. (2025); here g is exogenous.)

Figure 4 shows the trajectories. Under every growth model the human share collapses within the run: the end-state law of Section 3.2 in motion, h_t tracking e^*/e_t downward as e_t grows. The growth type sets the timing and the shape, not the destination — first-order growth produces a glide in which the human share halves with every capability doubling; the second-order term produces a cliff, since it reaches any capability level sooner than first-order growth alone (a speed claim, verified as such in the validation suite). Under the static model, coexistence at 0.44 was the end state; under measured improvement rates it is a transient.

3.4 Disposal of capital income (E2)

Figure 5a sweeps the disposal parameter r, the share of the owners' discretionary income spent back into the economy:

r	1.00	0.75	0.50	0.25	0.05	no AI
late output	27.0	21.4	17.2	15.6	19.7	24.6
AI share of wealth	0.51	0.52	0.52	0.48	0.27	—

Output ranges from above the no-AI reference (all income recycled as purchases) to well below it

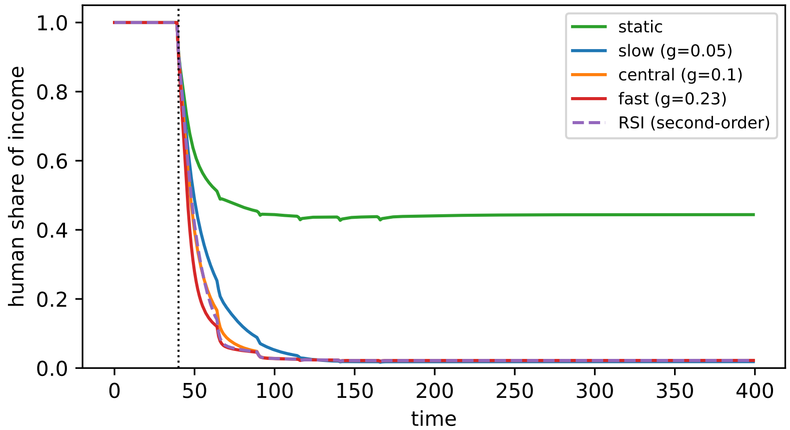

Figure 4: E6: human share of income under the capability-growth model types, g anchored to measured doubling-time ranges under the tick $\approx$ month reading (means over 8 seeds; dotted line: AI capital arrives).

(income retained; demand withdrawn), while the AI share of wealth is nearly constant over the entire upper range of r. The disposal of capital income determines the aggregate outcome; it does not determine the distributional one. The two statistics accordingly disagree most strongly at high r — precisely where output looks healthiest. One feature contradicted the pre-registered prediction and is reported as a finding: output has an interior minimum in r (19.7 at $r = 0.05$ against 15.6 at $r = 0.25$), because near-total retention also suppresses the capital's own upkeep and investment purchases, automation shares fall, and labor income partially recovers. The recovery reflects capital failing, not labor demand returning (Appendix B).

3.5 Interventions (E3)

Figure 5b sweeps the profit tax. From $\tau = 0$ to $\tau = 0.8$, household income rises from 10.8 to 16.7 (no-AI reference: 20.0) and the AI wealth share falls from 0.51 to 0.19, while output falls only from 27.0 to 25.5. The output cost of the tax is small because its revenue is recycled to households, whose spending replaces the reduced investment demand. At the default efficiency the tax thins the capital stock without eliminating it; per equation (2) the tax raises the survival threshold, and for efficiencies between $e^*(0)$ and $e^*(\tau)$ it eliminates the stock entirely — demonstrated directly in the validation suite, where capital that compounds untaxed at a marginal efficiency dies under the same-rate tax at the same efficiency.

Reassigning ownership is a different kind of lever. With title to a fraction $\omega = 0.4$ of each investment accruing to a public fund, the outcome is determined by the fund's disposal rule. A fund that distributes its earnings to households converges to exactly its diverted share (0.400) of the capital stock, with output 26.6 and household income 12.6: it acts as a transfer channel. A fund that reinvests on the same terms as private owners converges to full ownership (1.000) of the stock, with output 43.1 — the largest economy in this paper — and negligible private AI wealth: once per-unit profit settles at the fund's break-even point, private holdings contract by the factor $1 - \omega\delta$ per tick while the fund compounds (Figure 7, Appendix B). Neither design moves the survival threshold: ownership determines who benefits from the capital stock, not

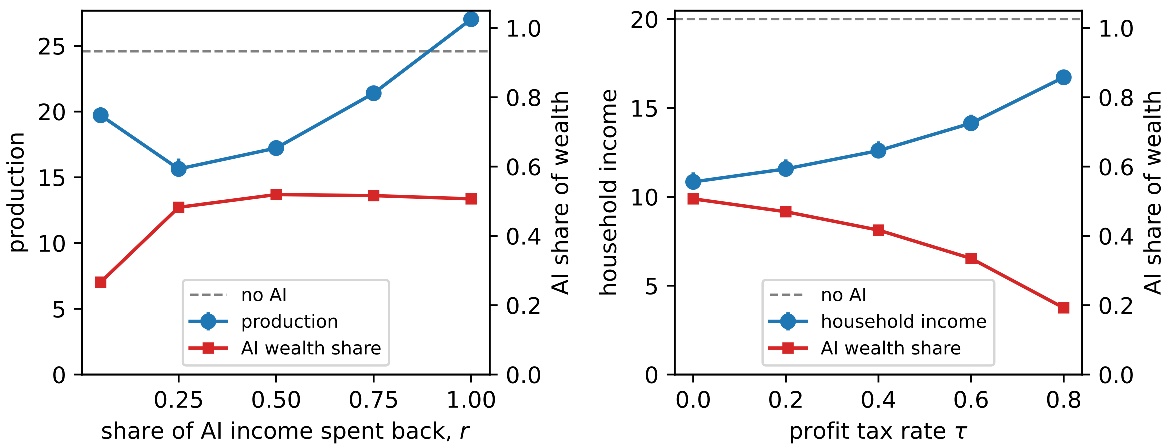

Figure 5: Disposal and taxation. (a) Sweeping the spent share r of AI income: late output (left axis) against the AI share of wealth (right axis). (b) Sweeping the profit tax τ: late household income (left axis) against the AI share of wealth (right axis). Error bars: bootstrap 95% intervals over 8 seeds.

whether it survives.

4 Discussion

Limitations. Five restrictions bound the scope of the results. First, no agent optimizes against the rules: all behavioral policies are fixed. The systems that would drive disempowerment in reality are adaptive; consequently every intervention that succeeds here succeeds against an opponent that does not respond, and each such result is an upper bound on effectiveness. An intervention that fails here fails under the weakest possible opposition, which makes the negative results the stronger ones. Second, there are no prices: the fixed-coefficient economy is the limiting case most favorable to human indispensability, since no substitution away from bottlenecked human input is possible. Third, there is no task creation: displaced labor is not reinstated in new activities, in contrast to Acemoglu and Restrepo (2018). Fourth, there is no financial system; investment is financed from retained profit only. Fifth — the assumptions that hold the human share up at moderate capability, made explicit by equation (3): human capacity neither exits nor is excluded (households always supply their unit of capacity and always receive the residual share; there is no unemployment margin), upkeep is linear in the capital stock, and, in the static baseline, capability is constant. Section 3.3 relaxes constancy directly and confirms the collapse; upkeep that falls with scale would act identically through the ratio e/m.

Conditions under which the conclusions change. Each mechanism identifies its own successor test. An adaptive capital owner that re-routes around the tax would break the threshold shift $s \mapsto s(1-\tau)$. Price adjustment would dissolve the fixed-coefficient bottleneck underlying the frozen-v threshold calculation. Upkeep costs that fall with scale would turn the fixed threshold into a moving one. Scale economies in the public fund would break the identical-technology condition behind the two fund regimes. Empirical evidence that the return-on-capacity of AI capital lies far above any plausible upkeep boundary would make the sub-threshold regime irrelevant in practice. Each of these is a defined, testable extension.

5 Conclusion

In a minimal conserved-money economy, the verbal argument of Kulveit et al. (2025) decomposes into one robust component and one conditional component. The robust component is distributional: once AI capital clears its survival threshold, the human share of income and wealth declines under every disposal rule and every intervention short of eliminating the capital stock. The conditional component is aggregate: whether output grows or contracts is decided entirely by where the capital's income goes, so aggregate output is uninformative about the distributional event — and most misleading exactly when it looks best. Among the interventions, taxation acts on the survival margin and can eliminate capital at small output cost in this closure; public ownership acts only on the assignment of claims. Extending the model with task creation, price adjustment, and adaptive owners — the three mechanisms most likely to overturn these orderings — is specified as future work, as are companion models of the cultural and political channels.

A Full specification

A.1 Motivation for the accounting discipline

Two earlier prototypes of this model produced qualitatively incorrect dynamics through violations of flow conservation: a household savings flow with no return path (aggregate income decayed geometrically before any AI capital arrived), and an AI consumption flow paid from no stock (unbounded output growth driven by demand injected by agents with zero wealth). Both defects were detected by a global conservation probe rather than by inspection of individual rules — if every flow rule conserves locally, their composition conserves, so a single global check suffices and any violation localizes to one rule by bisection. The specification therefore fixes the accounting first and derives the dynamics under it. One further correction is on record: the first version of the conserved quantity omitted the inventories-in-process term $\mathbf{1}^\top Ax$ and measured a spurious drift of +2.9% over 200 ticks; with the term included — as stock-flow-consistent practice requires (Godley and Lavoie, 2007) — the measured drift is float-precision rounding, and exactly zero without AI. The conserved stock is $M_t = (1-\sigma_s)\sum_h y_{h,t} + \sum W_t + \sum d_t^K + \mathbf{1}^\top Ax_t$; without AI systems the economy settles on a line of fixed points along which $\sum_h W_h^* / \sum_h y_h^* = \sigma_s/\sigma_d$ (the level of this ratio is an artifact of the tuned pair (σ_s, σ_d) and carries no economic content).

Money and title are separated exactly once: the AI pays the full investment purchase I (current account), while title to the resulting capital splits $(1-\omega)$ private, ω public (capital account). Sectors are pass-throughs with zero margin outside the capital share. The public fund's only income is its pro-rata capital revenue, its only outlays are upkeep and the dividend, and its stock grows only through ωI; its column therefore closes.

A.2 Dynamics

The model extends the register's Leontief substrate (`io_economy`: fixed coefficients, one Neumann step per tick, prices at unity; the dynamic reading of standard input-output analysis (Miller and Blair, 2009; Carvalho, 2014)) with owned capital stocks. AI system i holds $K_i \geq 0$ in a home sector; where the ownership-diversion mechanism is active, a public fund holds $K_j^{\text{pub}} \geq 0$ in each

Table 2: Transaction-flow matrix (one tick). Current-account rows are flows of money and each sums to zero across columns; capital-account rows record title to the capital purchased by the investment flow. + receives, − pays. Every symbol has a generating rule in Appendix A.2, eq. (4)–(5).

Current account	Households	Sectors	AI systems	Public fund
Household consumption C_H	−	+		
AI consumption $C_A = c_A W_A$		+	−	
Wages $(1-a_j)v_j x_j$	+	−		
Capital revenue, private share		−	+	
Capital revenue, public share		−		+
Upkeep paid, private $\min(\mathrm{rev}_i, mK_i)$		+	−	
Upkeep paid, public $\min(\mathrm{rev}^{\mathrm{pub}}, mK^{\mathrm{pub}})$		+		−
Investment purchase $I = \sum_i I_i$		+	−	
Profit tax $\tau\,[\pi_i]_+$	+		−	
Public dividend $d^{\mathrm{pub}} = \sum_j \left[\pi_j^{\mathrm{pub}}\right]_+$	+			−
Wealth accumulation				
$\Delta W_H = \sigma_s y_H - \sigma_d W_H$	✓			
$\Delta W_A = (1-s)\,[\pi]_+ - C_A$			✓	
Capital account (title to I)				
ΔK from investment			$+(1-\omega)I$	$+\omega I$

sector under identical technology. With sector capital $K_j^{\mathrm{tot}} = \sum_{i\in j} K_i + K_j^{\mathrm{pub}}$, the automation share is $a_j = eK_j^{\mathrm{tot}}/(eK_j^{\mathrm{tot}}+1)$. Capital revenue $a_j v_j x_j$ accrues to all owners in the sector pro rata by stock; upkeep mK is paid by every owner before profit; upkeep, investment, and AI consumption are all purchases from sectors. Per tick:

$$\pi_i = \mathrm{rev}_i - mK_i, \qquad \pi_j^{\mathrm{pub}} = \mathrm{rev}_j^{\mathrm{pub}} - mK_j^{\mathrm{pub}},$$

$$I_i = s\,[\pi_i]_+, \qquad C_{A,i} = c_A\,W_{A,i}, \tag{4}$$

$$K_i \leftarrow [(1-\delta)K_i + (1-\omega)I_i]_+, \qquad K_j^{\mathrm{pub}} \leftarrow \left[(1-\delta)K_j^{\mathrm{pub}} + \textstyle\sum_{i\in j}\omega I_i\right]_+, \tag{5}$$

with $W_{A,i} \leftarrow \left[W_{A,i} + (1-s)\,[\pi_i]_+ - C_{A,i}\right]_+$ the owner's retained stock. When $\pi_i \le 0$ the loss is charged to the capital stock ($K_i \leftarrow [K_i + \pi_i]_+$ before depreciation).

The survival threshold (2) follows from (5) directly: near $K = 0$ with v fixed, the stock evolves by the factor $1-\delta+s(ev-m)$, which exceeds one exactly when $e > (\delta/s+m)/v$. Under the profit tax the reinvested slice is $s(1-\tau)\,[\pi]_+$, giving $e^*(\tau)$ by the same algebra. The end-state share (3) follows from the steady-state balance: with $r = 1$ and no tax, reinvestment balances depreciation ($s\pi = \delta K$), so $a\,vx = (m+\delta/s)K$; substituting $K = a\,vx/(m+\delta/s)$ into $a = eK/(eK+1)$ gives $a^* = 1 - 1/\beta$ with $\beta = e\,vx/(m+\delta/s) = e/e^*$, hence $h^* = 1 - a^* = \min(1, e^*/e)$, where vx is the sector's value added in the automated steady state. E5's committed prediction evaluates vx at its people-only level; since automation raises value added, the prediction is an upper bound on the human share. The contraction factor $1-\omega\delta$ of the reinvesting public fund follows from the common per-unit profit ρ_t: at the fund's break-even $\rho^* = \delta/s$, the private update factor is $(1-\delta) + (1-\omega)s\rho^* = 1 - \omega\delta$.

Upkeep settlement. Upkeep is paid out of revenue only: the sector receives $\min(\mathrm{rev}, mK)$ from each owner. The unpaid remainder is not a monetary flow; it is physical deterioration charged to the stock by the loss rule. No agent pays money it does not hold, and the min appears explicitly in Table 2, so the matrix closes in the loss case.

Disposal family. The parameter $r \in [0, 1]$ scales the owners' discretionary outlays: investment purchases become rI_i (capital formation scaled identically) and consumption becomes $rc_A W_{A,i}$; the unspent remainder accumulates in $W_{A,i}$. Money is conserved at every r: retention is a demand stall — value held in a recorded stock with zero velocity — not a leak. The baseline equations are the $r = 1$ member.

Households and the public fund. Households receive the labor share $(1 - a_j)v_j x_j$, spend $(1 - \sigma_s)$ of income plus the drawdown $\sigma_d W_H$, and receive tax revenue and public dividends as equal transfers where those mechanisms are active. In the distributing design the fund pays $d^{\text{pub}} = \sum_j \left[\pi_j^{\text{pub}}\right]_+$ to households and grows only through ωI; in the reinvesting design it retains the same s-slice as private owners.

A.3 Reference substrates

Two register members serve as structural comparisons. S1 (`compute_economy`) is a one-equation CES economy with marginal-product factor payments; it is retained only as a contrast case, since for substitution elasticities above unity the decline of the human share is implied by the functional form rather than found. S2 (`io_economy`) is the plain Leontief economy without capital stocks; its independently measured endpoints (activity $21.3 \to 32.7$ with human share $\to 0.01$ at full reinvestment; activity $\to 0$ at recycling 0.3) and an independent JavaScript implementation of the same dynamics provide a cross-implementation check that the disposal mechanism is not an artifact of a single codebase.

B Experimental details

Predictions committed before any sweep ran: (i) E1 exhibits both regimes, with the empirical threshold inside the analytic e^* band evaluated on measured sector value added; (ii) E2 shows the ownership share rising for every r and output switching from non-decreasing to contracting below a closure threshold; (iii) E3 shows the elimination band on the τ axis and no threshold shift on the ω axis (the latter as an implementation-fidelity check). All sweeps: 8 shared seeds, 400 ticks, bootstrap confidence intervals, means over the final quarter of each run.

E1 (survival threshold). Grid over (e, s); both regimes are required by design — a sweep that never exhibits the elimination regime would be uninformative about the threshold. Figure 6: at $s = 0.5$, late private capital is 0.00 for $e \leq 0.2$, ignites at $e = 0.3$ (the hub sector, whose higher value added yields the lowest threshold; the pre-registered band for ignition was $[0.26, 0.70]$), and increases through the band above. The prediction held at both reinvestment rates.

E2 (disposal). Reported in Section 3.4. The interior minimum of output in r contradicted prediction (ii) and is recorded as a finding: at $r = 0.05$ the AI share of wealth is falling because the capital stock is failing — retained income suppresses the capital's own upkeep and investment purchases — not because labor demand recovered. Prediction (ii)'s qualitative content (share rises for every r over the compounding phase; output contracts below a closure threshold) held.

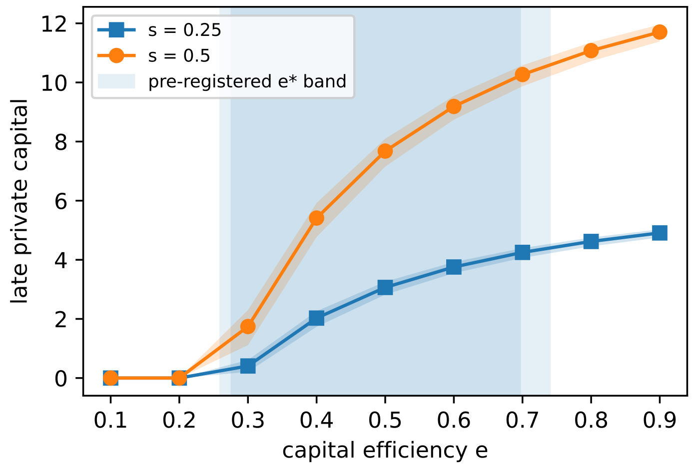

Figure 6: E1: late private capital against efficiency e at two reinvestment rates. Shaded bands: pre-registered analytic e^* ranges, evaluated on measured sector value added and committed before the sweep ran.

E5 (capability and the end state). Sweep of e over $[0.1, 16]$, log-spaced, at otherwise default parameters. The prediction $h^* = \min(1, e^*/e)$ — evaluated per sector on people-only value added and aggregated with value-added weights — was committed before the sweep ran. The measured shares lie at or below the prediction throughout the compounding regime (0.444 measured against 0.792 predicted at the baseline $e = 0.55$) and converge to it at high capability (0.032 against 0.030 at $e = 16$). The gap has a single direction and a single cause: the prediction freezes value added at its no-AI level, while automation raises value added and shifts composition toward automated sectors. Reported as measured in Figure 3.

E6 (capability growth). The five growth models of Section 3.3 at otherwise default parameters. Late human shares: 0.444 static; 0.022 for every growing model (all reach the numerical capability cap $e_{\max} = 64$ within the run, at times ordered by growth type). The validation suite verifies the mechanism claims: with $g = \gamma = 0$ the traced capability is exactly constant (the growth law is limit-equivalent to the static model), the second-order term crosses any capability level sooner than first-order growth alone, and money remains conserved under growth — a moving e only reallocates value-added shares. Note that at matched horizons past the cap, the shares of the growing models need not be ordered by growth speed; the second-order claim is about timing, not depth.

E3 (interventions). Reported in Section 3.5. The elimination band is demonstrated in the validation suite at the hub sector's marginal efficiency. The ownership arm at $\omega = 0.4$ reproduces both fund regimes (Figure 7). One correction is on record in accordance with the pre-registration discipline: an earlier draft claimed the ownership diversion output-neutral by an additivity argument requiring only identical earning rates; the implementation falsified this under both disposal designs (distributing: aggregate capital -18%; reinvesting: public share 0.999, output $+61\%$, single seed at $\omega = 0.4$), because additivity requires identical *disposal*, and any payout difference changes the demand path. The two regimes reported in Section 3.5 are the corrected

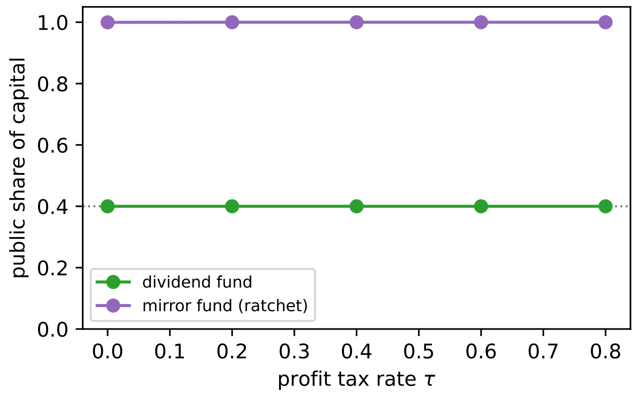

Figure 7: E3, ownership arm: public share of the capital stock under the two fund designs. The distributing fund holds its diverted share ω exactly; the reinvesting fund converges to full ownership.

result.

E4 (structural robustness). The register discipline requires re-running the E1/E3 orderings on the reference substrates of Appendix A.3 where defined. Currently reported: S2's independently measured decoupling endpoints and the cross-implementation replication; S1's closed-form tax threshold as the analytic cross-reference for the threshold shift. The full cross-substrate rerun is future work.

References

Daron Acemoglu. The simple macroeconomics of AI. *Economic Policy*, 40(121):13–58, 2025. doi: 10.1093/epolic/eiae042. Also NBER Working Paper 32487 (2024).

Daron Acemoglu and Pascual Restrepo. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6): 1488–1542, 2018. doi: 10.1257/aer.20160696.

Philippe Aghion, Benjamin F. Jones, and Charles I. Jones. Artificial intelligence and economic growth. In Ajay Agrawal, Joshua Gans, and Avi Goldfarb, editors, *The Economics of Artificial Intelligence: An Agenda*, pages 237–282. University of Chicago Press, Chicago, 2019. Also NBER Working Paper 23928 (2017).

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–157, 2011. doi: 10.3982/ECTA8416.

Erik Brynjolfsson. The Turing trap: The promise and peril of human-like artificial intelligence. *Daedalus*, 151(2):272–287, 2022. doi: 10.1162/daed_a_01915.

Erik Brynjolfsson, Daniel Rock, and Chad Syverson. Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In Ajay Agrawal, Joshua Gans, and Avi Goldfarb, editors, *The Economics of Artificial Intelligence: An Agenda*, pages 23–57. University of Chicago Press, Chicago, 2019. Also NBER Working Paper 24001 (2017).

Vasco M. Carvalho. From micro to macro via production networks. *Journal of Economic Perspectives*, 28(4):23–48, 2014. doi: 10.1257/jep.28.4.23.

Ege Erdil, Andrei Potlogea, Tamay Besiroglu, Edu Roldan, Anson Ho, Jaime Sevilla, Matthew Barnett, Matej Vrzla, and Robert Sandler. GATE: An integrated assessment model for AI automation, 2025.

J. Doyne Farmer and Duncan Foley. The economy needs agent-based modelling. *Nature*, 460 (7256):685–686, 2009. doi: 10.1038/460685a.

Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1:255–294, 2009. doi: 10.1146/annurev.economics.050708.142940.

Wynne Godley and Marc Lavoie. *Monetary Economics: An Integrated Approach to Credit, Money, Income, Production and Wealth.* Palgrave Macmillan, Basingstoke, UK, 2007. ISBN 9780230500556.

Anson Ho, Tamay Besiroglu, Ege Erdil, David Owen, Robi Rahman, Zifan Carl Guo, David Atkinson, Neil Thompson, and Jaime Sevilla. Algorithmic progress in language models, 2024.

Anton Korinek and Joseph E. Stiglitz. Artificial intelligence and its implications for income distribution and unemployment. In Ajay Agrawal, Joshua Gans, and Avi Goldfarb, editors, *The Economics of Artificial Intelligence: An Agenda*, pages 349–390. University of Chicago Press, Chicago, 2019. Also NBER Working Paper 24174 (2017).

Jan Kulveit, Raymond Douglas, Nora Ammann, Deger Turan, David Krueger, and David Duvenaud. Gradual disempowerment: Systemic existential risks from incremental AI development, 2025.

Thomas Kwa, Ben West, Joel Becker, Amy Deng, Katharyn Garcia, Max Hasin, et al. Measuring ai ability to complete long software tasks, 2025. METR.

Ronald E. Miller and Peter D. Blair. *Input-Output Analysis: Foundations and Extensions*. Cambridge University Press, Cambridge, UK, 2 edition, 2009. ISBN 9780521517133.

Pascal Seppecher, Isabelle Salle, and Marc Lavoie. What drives markups? evolutionary pricing in an agent-based stock-flow consistent macroeconomic model. *Industrial and Corporate Change*, 2018. doi: 10.1093/ICC/DTY011.

Jaime Sevilla, Lennart Heim, Anson Ho, Tamay Besiroglu, Marius Hobbhahn, and Pablo Villalobos. Compute trends across three eras of machine learning. In *2022 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022.

Joseph Zeira. Workers, machines, and economic growth. *The Quarterly Journal of Economics*, 113(4):1091–1117, 1998. doi: 10.1162/003355398555847.