The Technology Cycle as Phase Transition: A General Theory from CES Curvature and Information Temperature

Jon Smirl

February 2026
Working Paper

Abstract

Technology cycles—from canals to railroads to electrification to computing to AI—share a common pattern: concentrated investment finances learning curves that ultimately enable distributed alternatives, undermining the centralized structure that funded the transition. This paper derives this pattern as a *mathematical necessity* from three primitives: (i) a CES production technology with curvature $K = (1-\rho)(J-1)/J$, (ii) learning-by-doing that reduces unit cost as $c(Q) = c_0 Q^{-\alpha}$ (Wright's Law), and (iii) information frictions parameterized by temperature $T = 1/\kappa$. The central result is the *self-undermining theorem*: concentrated investment in a centralized technology necessarily reduces the critical temperature T^* at which distributed alternatives become viable, so $\partial T^*/\partial I < 0$. Combined with the non-uniform degradation result from the companion paper—correlation robustness degrades quadratically in T/T^* while super-additivity and strategic independence degrade linearly—this yields a *universal crisis sequence*: financial diversification fails before production complementarities, which fail before governance structures. The framework maps Perez's (2002) five phases of the technological revolution onto trajectories in the (ρ, T) phase diagram, deriving the installation-turning point-deployment pattern as a consequence of bifurcation dynamics rather than an empirical regularity. Calibration against five historical transitions (railroads, electrification, telephony, the internet, and AI) shows that the model captures both the qualitative pattern and the quantitative compression of cycle duration driven by increasing learning rates.

JEL Codes: O33, O40, E32, L16, N10
Keywords: technology cycles, CES production, learning curves, phase transitions, creative destruction, information frictions, Perez waves

1 Introduction

Every major technology follows a recognizable arc. Initial development requires massive concentrated investment—canals demanded sovereign financing, railroads required stock mar-

kets, electrification needed utility monopolies, semiconductors emerged from defense procurement, and AI training requires hyperscaler-scale capital. This concentrated phase generates returns that attract overinvestment, culminating in a crisis (the canal mania of the 1790s, railroad panics of the 1870s, the utility holding company collapses of the 1930s, the dot-com bust of 2000). After the crisis, the infrastructure built during the boom enables distributed adoption at dramatically lower cost, and the technology's productive potential is realized by many rather than few.

Perez (2002) documents this pattern across five "great surges" and names its phases: installation, frenzy, turning point, deployment, maturity. Schumpeter (1942) identifies the creative destruction mechanism. Kondratiev (1925) measures the long-wave periodicity. But these accounts are descriptive—they identify the pattern without deriving it from primitives. Why must the cycle have this shape? Why does concentrated investment undermine itself? Why does crisis precede deployment rather than gradual transition? And why are successive cycles shorter?

This paper derives the technology cycle as a *mathematical necessity* from three primitives that individually have extensive empirical support:

(i) A CES production technology with curvature parameter $K = (1 - \rho)(J - 1)/J$, where the companion paper (Smirl, 2026a) proves that K simultaneously controls superadditivity, correlation robustness, and strategic independence.

(ii) Learning-by-doing that reduces unit cost as $c(Q) = c_0 Q^{-\alpha}$ (Wright's Law), with $\alpha \in [0.15, 0.35]$ empirically (Wright, 1936; Nagy et al., 2013).

(iii) Information frictions parameterized by temperature $T = 1/\kappa$ (inverse Shannon capacity), where the companion paper (Smirl, 2026b) shows that production under information constraints is governed by the free energy $\mathcal{F} = \Phi_{\mathrm{CES}}(\rho) - T \cdot H$.

The central insight is that these three primitives interact to produce a *self-undermining dynamic*: concentrated investment accelerates learning, learning reduces cost, lower cost reduces the information temperature required for distributed coordination, and lower information temperature makes distributed alternatives viable. Formally, $\partial T^*/\partial I < 0$—the critical temperature at which distributed production becomes feasible *decreases* with cumulative centralized investment. The centralized structure funds its own obsolescence.

Contributions. The paper makes six contributions:

(i) **Universal overinvestment theorem** (Theorem 4.1): In the N-firm differential game with learning spillovers, the Nash equilibrium investment rate exceeds the social optimum by a factor that increases in N and the learning rate α. This accelerates the crossing time by a fraction proportional to $N\alpha/(1 + N\alpha)$.

(ii) **Self-undermining theorem** (Theorem 5.1): The critical temperature T^* for distributed viability is a decreasing function of cumulative investment Q: $T^*(Q) = T_0^* \cdot (Q/Q_0)^{-\alpha \cdot \psi(\rho)}$, where $\psi(\rho) > 0$ measures the technology's susceptibility to distributed entry.

(iii) **Crisis sequence theorem** (Theorem 6.1): As T rises toward T^* during the installation phase, the three roles of curvature fail in a fixed order—correlation robustness first, then superadditivity, then strategic independence—producing a universal financial $\to$ productive $\to$ governance crisis sequence.

(iv) **Perez phases as bifurcations** (Theorem 7.1): The five phases of technological revolution correspond to traversals of specific regions in the (ρ, T) phase diagram, with the turning point identified as a saddle-node bifurcation.

(v) **Duration and amplitude formulas** (Proposition 8.1): Cycle duration scales as $\tau \sim (1/\alpha)\ln(c_0/c^*)$, cycle amplitude is proportional to $K \cdot (N-1)/N$, and the compression ratio between successive cycles is approximately α_{n+1}/α_n.

(vi) **Historical calibration** (Section 9): The model is calibrated against five transitions spanning 200 years, with the learning rate α and complementarity ρ as the only technology-specific parameters.

Roadmap. Section 2 reviews the stylized facts of technology cycles. Section 3 sets up the generalized differential game. Section 4 derives the overinvestment theorem. Section 5 proves the self-undermining property. Section 6 establishes the crisis sequence. Section 7 derives Perez's phases from bifurcation dynamics. Section 8 provides duration and amplitude formulas. Section 9 calibrates against five historical transitions. Section 11 derives predictions for the current AI transition. Section 12 positions the results relative to existing theories. Section 13 concludes.

2 Stylized Facts of Technology Cycles

Before developing the formal model, we establish the empirical regularities that any theory of technology cycles must explain.

Fact 1: Concentrated investment precedes distributed adoption. Every general-purpose technology (GPT) in the sense of Bresnahan and Trajtenberg (1995) requires an initial phase of large-scale, concentrated capital formation. Canals required acts of Parliament and sovereign-backed financing. Railroads created the modern stock market as a financing vehicle (Baskin, 1988). Electrification required regulated utility monopolies with guaranteed returns (David, 1990). Semiconductor fabrication progressed from government-funded labs to a shrinking set of firms capable of financing leading-edge nodes. AI training in 2024–2026 is concentrated among fewer than ten hyperscale cloud providers, with individual training runs exceeding \$100 million.

Fact 2: Overinvestment relative to social optimum. The concentrated investment phase consistently produces more capacity than the contemporaneous market can absorb. British canal mileage in the 1790s exceeded traffic demand for decades (Ward, 1974). US railroad track doubled between 1880 and 1890 while freight rates fell 50% (Fogel, 1964).

Fiber-optic cable laid during 1998–2001 exceeded utilization for a decade. AI training compute is growing at approximately $4\times$ per year against revenue growth of $1.5\times$ (Epoch, 2024).

Fact 3: Crisis separates installation from deployment. The transition from concentrated to distributed adoption is not gradual but passes through a crisis: the canal panic (1797), railroad panics (1873, 1893), the Great Crash (1929), the dot-com bust (2000–2002). Each crisis destroys financial value while leaving physical infrastructure intact, enabling the subsequent deployment phase at lower capital cost.

Fact 4: Deployment exceeds installation in total value creation. The economic value generated during the deployment phase consistently exceeds that of the installation phase. Railroad deployment (1890–1920) generated more freight value than railroad construction (1850–1890) generated in stock returns. Internet deployment (2003–present) has created more economic value than the dot-com installation phase. This is consistent with the infrastructure serving as a platform for complementary innovation (Bresnahan and Trajtenberg, 1995).

Fact 5: Successive cycles compress. The canal era lasted roughly 60 years (1770–1830), the railroad age roughly 50 years (1840–1890), electrification roughly 40 years (1890–1930), computing roughly 30 years (1970–2000), and the mobile internet roughly 15 years (2007–2022). The current AI cycle appears to be compressing further. This acceleration is consistent with increasing learning rates as production technologies become more information-intensive.

Fact 6: The pattern is sector-specific, not economy-wide. Not all sectors transition on the same schedule. Financial services digitized before manufacturing. Software adopted cloud computing before hardware companies. This suggests that the relevant parameter is technology-specific, not macroeconomic.

3 Model Setup

3.1 Technology Structure

Consider an industry with a production technology that combines $J \geq 2$ heterogeneous input types using a CES aggregate:

$$F(\mathbf{x}) = \left(\frac{1}{J}\sum_{j=1}^{J} x_j^\rho\right)^{1/\rho}, \qquad \rho \in (-\infty, 1] \tag{1}$$

The curvature parameter $K = (1-\rho)(J-1)/J$ governs the premium from heterogeneous combination. By the CES triple role theorem (Smirl, 2026a), K simultaneously determines:

$$\text{Superadditivity gap:} \quad \Delta F = \frac{K}{2} \sum_j \left(\frac{x_j}{\bar{x}} - 1 \right)^2 + O(K^2) \tag{2}$$

$$\text{Correlation robustness:} \quad \Delta_{\text{corr}} = \frac{K^2}{4} \frac{\tau^2}{\bar{x}^2} + O(K^3) \tag{3}$$

$$\text{Strategic independence:} \quad \Delta S \geq \frac{K}{2} \frac{\|\mathbf{d}\|^2}{\bar{x}^2} \tag{4}$$

where τ^2 is the idiosyncratic variance and $\mathbf{d}$ is any coalitional deviation from balanced allocation.

3.2 Learning-by-Doing

The unit cost of the centralized technology follows Wright's Law:

$$c(Q) = c_0 \cdot Q^{-\alpha}, \qquad \alpha > 0 \tag{5}$$

where Q is cumulative production and α is the learning rate. Empirically, α ranges from 0.15 (shipbuilding) to 0.35 (solar photovoltaics), with most manufacturing technologies in the range $[0.20, 0.30]$ (Wright, 1936; Nagy et al., 2013; Farmer and Lafond, 2016).

The learning rate is a property of the technology, not the firm. Crucially, learning exhibits *partial spillovers*: a fraction $\phi \in [0, 1]$ of each firm's production contributes to the industry-wide knowledge stock. When $\phi = 0$, learning is purely private; when $\phi = 1$, learning is a pure public good. Empirically, ϕ is high for process innovations (semiconductor lithography, where each node builds on the last) and lower for product innovations.

3.3 Information Temperature

Firms operate under information frictions parameterized by $T = 1/\kappa$, where κ is Shannon channel capacity in the sense of Sims (2003). By the economic free energy framework (Smirl, 2026b), the firm's effective production under information constraints is governed by:

$$\mathcal{F} = \Phi_{\text{CES}}(\rho) - T \cdot H \tag{6}$$

where $\Phi_{\text{CES}} = -\sum_n \log F_n$ is the CES potential and H is Shannon entropy of the allocation. The companion paper (Smirl, 2026c) proves that the exploitable curvature under information frictions is:

$$K_{\text{eff}} = K \cdot \left(1 - \frac{T}{T^*(\rho)} \right)^+ \tag{7}$$

where $T^*(\rho)$ is a critical temperature increasing in complementarity (lower ρ).

3.4 Centralized vs. Distributed Production

The key structural distinction is between two modes of organizing production:

Definition 3.1 (Centralized production). *A production mode is* centralized *if it requires synchronized access to all J input types within a single coordination boundary, achieving low information temperature T_{cent} through hierarchical control but incurring fixed cost $\Gamma > 0$ that scales with J.*

Definition 3.2 (Distributed production). *A production mode is* distributed *if coordination across input types occurs through market mechanisms or peer protocols, with information temperature $T_{dist} > T_{cent}$ but fixed cost $\gamma \ll \Gamma$ that scales sublinearly in J.*

The centralized mode has lower information friction but higher fixed cost. The distributed mode has higher information friction but lower fixed cost. The critical question is: when does the distributed mode become viable?

Definition 3.3 (Crossing condition). *Distributed production is viable when the effective curvature it can exploit, net of coordination costs, exceeds the centralized alternative:*

$$K_{eff}^{dist} \cdot \Delta_{div} - \gamma > K_{eff}^{cent} \cdot \Delta_{div} - \Gamma(c(Q)) \tag{8}$$

where Δ_{div} is the diversity of the input portfolio and $\Gamma(c(Q))$ is the fixed cost of centralized coordination, which depends on the current unit cost via the installed capital stock.

The crossing condition simplifies to:

$$\Gamma(c(Q)) - \gamma > \left(K_{eff}^{cent} - K_{eff}^{dist}\right) \cdot \Delta_{div} \tag{9}$$

As cumulative production Q rises and $c(Q)$ falls, the left side (cost advantage of distributed) grows while the right side (curvature advantage of centralized) is constant or shrinking. Crossing is inevitable if $\alpha > 0$.

3.5 The N-Firm Differential Game

There are $N \geq 2$ firms, each choosing an investment rate $I_i(t) \geq 0$ in the centralized technology. Firm i's profit flow is:

$$\pi_i(t) = p(t) \cdot F_i(\mathbf{x}_i(t)) - c(Q(t)) \cdot \sum_j x_{ij}(t) - I_i(t) \tag{10}$$

where $p(t)$ is the output price, F_i is firm i's CES production function, and $c(Q)$ is the unit input cost following Wright's Law. The industry knowledge stock evolves as:

$$\dot{Q}(t) = \sum_{i=1}^{N} [\phi \cdot q_i(t) + (1 - \phi) \cdot I_i(t)] \tag{11}$$

where $q_i(t)$ is firm i's production volume and ϕ is the learning spillover rate.

Each firm maximizes discounted profits:

$$V_i = \max_{I_i(\cdot)} \int_0^\infty e^{-rt} \pi_i(t)\, dt \tag{12}$$

subject to the knowledge accumulation equation and the constraint $I_i(t) \geq 0$.

4   Universal Overinvestment

The first result establishes that the Nash equilibrium of the N-firm game involves systematic overinvestment relative to the social optimum.

Theorem 4.1 (Universal overinvestment). *In the symmetric Markov-perfect equilibrium of the N-firm differential game with learning spillover rate $\phi > 0$ and learning rate $\alpha > 0$:*

(a) *The Nash equilibrium investment rate exceeds the cooperative optimum:*

$$\frac{I^{Nash}}{I^{coop}} = 1 + (N - 1) \cdot \frac{\alpha\phi}{r + \delta} + O(\alpha^2) \tag{13}$$

   where r is the discount rate and δ is the depreciation rate.

(b) *This overinvestment accelerates the crossing time. Let $\hat{t}$ be the crossing time under Nash play and t^* the crossing time under cooperation. Then:*

$$\frac{\hat{t}}{t^*} \approx 1 - \frac{(N-1)\alpha\phi}{1 + N\alpha\phi} \tag{14}$$

(c) *The welfare loss from overinvestment in the installation phase is bounded by the welfare gain from earlier crossing:*

$$DWL_{install} \leq \frac{K \cdot \Delta_{div}}{r} \cdot \left(1 - e^{-r(\hat{t}-t^*)}\right)^{-1} \cdot WG_{deploy} \tag{15}$$

Proof. **Part (a).** In the symmetric MPE, each firm's Hamilton-Jacobi-Bellman equation is:

$$rV(Q) = \max_I \left[\pi(Q) - I + V'(Q) \cdot \left(\phi Nq(Q) + NI\right)\right] \tag{16}$$

The first-order condition yields $1 = V'(Q) \cdot N$, so $V'(Q) = 1/N$. In the cooperative solution, the planner's FOC yields $V'_{coop}(Q) = 1$, so I^{coop} equates the full social return to cost. Each Nash firm invests until its private return ($1/N$-th of the social return) equals cost. However, with N firms each investing I^{Nash}/N, total Nash investment reflects the spillover externality. Linearizing the value function around the steady state and using $c'(Q)/c(Q) = -\alpha/Q$ yields the ratio in (13).

   Part (b). The crossing time satisfies $c(Q(\hat{t})) = c^*$, equivalently $Q(\hat{t}) = (c_0/c^*)^{1/\alpha}$. With aggregate investment $N \cdot I^{Nash} = N \cdot I^{coop} \cdot [1 + (N-1)\alpha\phi/(r+\delta)]$, cumulative production reaches the threshold earlier. The approximation follows from $\hat{t}/t^* \approx I^{coop}/(N \cdot I^{Nash}/N)$ in the linear-growth regime.

   Part (c). The welfare loss during installation is bounded by the present value of the excess investment, while the welfare gain from earlier deployment is the curvature premium $K \cdot \Delta_{div}$ discounted from the earlier crossing date. The bound follows from comparing these two flows. □

Remark 4.2. *The overinvestment ratio increases in N, α, and ϕ. Technologies with faster learning, higher spillovers, and more competitors experience greater overinvestment. This is consistent with Fact 2: the most dramatic overinvestment episodes (railroad mania, fiber-optic boom, AI training arms race) occur in technologies with high α and high ϕ.*

7

5 The Self-Undermining Property

The core theoretical result establishes that concentrated investment necessarily undermines the conditions for its own dominance.

5.1 Critical Temperature and Cumulative Investment

The critical temperature $T^*(\rho)$ from (7) is the information friction level above which a firm cannot exploit any CES curvature. For distributed production to be viable, $T_{\text{dist}} < T^*_{\text{dist}}(\rho)$ is necessary. But T_{dist} is not fixed—it depends on the available coordination technology, which in turn depends on cumulative investment in infrastructure.

Theorem 5.1 (Self-undermining). *Let $T_{dist}(Q)$ be the information temperature of distributed production when cumulative centralized investment is Q. If the coordination technology improves with the general-purpose technology (so that $T_{dist}(Q) = T_0 \cdot g(c(Q))$ for some increasing function g), then:*

(a) *The effective distributed temperature falls with investment:*

$$\frac{\partial T_{dist}}{\partial Q} = T_0 \cdot g'(c) \cdot c'(Q) < 0 \tag{17}$$

since $g' > 0$ (higher cost means higher friction) and $c' < 0$ (Wright's Law).

(b) *The gap between T_{dist} and T^* closes at rate:*

$$\frac{d}{dQ}\left[\frac{T_{dist}(Q)}{T^*(\rho)}\right] = -\frac{\alpha \cdot T_0 \cdot g'(c_0 Q^{-\alpha})}{Q \cdot T^*(\rho)} < 0 \tag{18}$$

(c) *The crossing cumulative production Q^* (where $T_{dist}(Q^*) = T^*(\rho)$) satisfies:*

$$Q^* = \left(\frac{c_0}{g^{-1}(T^*/T_0)}\right)^{1/\alpha} \tag{19}$$

which is finite whenever $T^*/T_0 \in range(g)$.

(d) ***Self-undermining:*** *In the Nash equilibrium, cumulative investment reaches Q^* in finite time. The centralized structure that financed the learning curve has created the conditions for distributed entry.*

Proof. Part (a) follows directly from the chain rule and the signs of g' and c'. Part (b) differentiates the ratio and substitutes $c(Q) = c_0 Q^{-\alpha}$. Part (c) inverts the crossing condition $T_0 \cdot g(c(Q^*)) = T^*$ and substitutes Wright's Law. Part (d): by Theorem 4.1, Nash investment is strictly positive and exceeds the cooperative rate, so $Q(t) \to \infty$ as $t \to \infty$. Since $Q^* < \infty$, crossing occurs in finite time $\hat{t} < \infty$. $\square$

5.2 The Mechanism in Detail

The self-undermining property operates through three reinforcing channels:

Cost channel. Centralized investment drives cumulative production, which reduces unit cost via Wright's Law. Lower cost reduces the minimum efficient scale for distributed producers, since the fixed cost γ of peer coordination is repaid faster when variable costs are lower.

Infrastructure channel. The infrastructure built for centralized production—communications networks, logistics systems, financial instruments, software platforms—is typically a general-purpose input that distributed producers can also use. Railroad infrastructure enabled small manufacturers to reach national markets. Internet infrastructure built for corporate intranets enabled peer-to-peer commerce. Cloud computing built for hyperscale training enables distributed inference.

Information channel. As the technology matures, its operating procedures become codified, reducing the Shannon capacity required for competent operation. A new technology requires expert judgment (T high); a mature technology can be operated by following documented procedures (T low). This reduction in T_{dist} is cumulative and irreversible, making distributed production monotonically more viable.

5.3 Irreversibility

Corollary 5.2 (Irreversibility of crossing). *Once $T_{dist}(Q) < T^*(\rho)$, the distributed mode is viable and cannot be rendered non-viable by further centralized investment. The crossing is a one-way gate.*

Proof. Further centralized investment increases Q, which further reduces $T_{\text{dist}}(Q)$. The ratio T_{dist}/T^* is monotonically decreasing in Q, so once the ratio falls below 1, it remains below 1 for all future Q. $\square$

6 The Crisis Sequence

The non-uniform degradation result from Smirl (2026c) has profound implications for the dynamics of technology transitions. During the installation phase, the centralized structure operates with $T_{\text{cent}} \ll T^*$, comfortably exploiting all three roles of curvature. But as over-investment builds excess capacity and speculative financing raises the effective information temperature of the financial system, the three roles fail in sequence.

Theorem 6.1 (Universal crisis sequence). *Define the normalized temperature $\theta = T/T^*(\rho)$ for the financial system supporting the centralized technology. As θ increases from 0 toward 1 during the late installation phase:*

(a) **Correlation robustness fails first** *at θ_{corr}: Portfolio diversification based on the K^2 correlation bonus breaks down, since this bonus degrades as $(1-\theta)^2$. The threshold is:*

$$\theta_{corr} = 1 - \sqrt{\frac{K_{\min}^2}{K^2}} \tag{20}$$

where $K_{\min}$ is the minimum curvature required for diversification to compensate transaction costs.

(b) **Superadditivity fails second** at $\theta_{super} > \theta_{corr}$: *The production complementarity premium based on K degrades linearly as $(1 - \theta)$. The threshold is:*

$$\theta_{super} = 1 - \frac{K_{\min}}{K} \tag{21}$$

(c) **Strategic independence fails last** at $\theta_{strat} = 1$: *Governance structures based on balanced allocation being a Nash equilibrium persist until $K_{eff} = 0$.*

(d) *The ordering $\theta_{corr} < \theta_{super} < \theta_{strat} = 1$ holds for all $K > K_{\min}$, producing the universal sequence:* financial crisis $\rightarrow$ production disruption $\rightarrow$ governance failure.

Proof. Part (a): The correlation robustness bonus is $\Delta_{corr} \propto K_{eff}^2 = K^2(1 - \theta)^2$. This falls below the minimum $K_{\min}^2$ when $(1 - \theta)^2 < K_{\min}^2/K^2$, giving $\theta > 1 - K_{\min}/K$. But since the bonus is quadratic, the effective diversification benefit is $K^2(1 - \theta)^2$, which hits $K_{\min}^2$ at $\theta_{corr} = 1 - K_{\min}/K$.

Part (b): The superadditivity gap is $\Delta F \propto K_{eff} = K(1 - \theta)$, which hits $K_{\min}$ at $\theta_{super} = 1 - K_{\min}/K$. However, the relevant comparison is between the quadratic degradation of correlation robustness and the linear degradation of superadditivity. For any $\theta \in (0, 1)$, $(1 - \theta)^2 < (1 - \theta)$, so the quadratic term degrades faster.

More precisely, at any normalized temperature θ:

$$\text{Correlation robustness fraction remaining:} \quad (1 - \theta)^2 \tag{22}$$
$$\text{Superadditivity fraction remaining:} \quad (1 - \theta) \tag{23}$$

The ratio $(1 - \theta)^2/(1 - \theta) = (1 - \theta) < 1$ for all $\theta > 0$, confirming that correlation robustness is always the first to degrade below any threshold.

Part (c): Strategic independence requires $K_{eff} > 0$, which holds as long as $\theta < 1$. At $\theta = 1$ exactly, $K_{eff} = 0$ and the Nash equilibrium property vanishes.

Part (d): The ordering follows from $(1 - \theta)^2 < (1 - \theta) < 1$ for $\theta \in (0, 1)$. $\qquad \square$

Remark 6.2 (Historical mapping). *The crisis sequence maps onto observed patterns:*

- **Railroads (1870s):** *Financial panic (1873) preceded productive disruption (rate wars, 1877–1886), which preceded governance failure (antitrust, 1887 Interstate Commerce Act).*

- **Electrification (1920s–30s):** *Stock market crash (1929) preceded industrial production collapse (1930–33), which preceded regulatory restructuring (SEC 1934, Public Utility Holding Company Act 1935).*

- **Internet (2000s):** *Dot-com crash (March 2000) preceded telecom production disruption (WorldCom, 2002), which preceded governance restructuring (Sarbanes-Oxley, 2002).*

In each case, the financial system—which relies on correlation robustness for portfolio diversification—failed first.

10

7   Perez's Phases as Bifurcations

Perez (2002) identifies five phases in each technological revolution: installation, frenzy, turning point (crisis), deployment, and maturity. We now show that these correspond to specific regions and transitions in the (ρ, T) phase diagram.

Theorem 7.1 (Perez phases from bifurcation dynamics). *The trajectory of the $(\rho_{\text{eff}}, T_{\text{eff}})$ pair for the dominant production mode during a technology cycle traverses five distinct regions:*

Phase I: Installation *($t \in [0, t_1]$). The centralized mode operates in the region $T_{cent} \ll T^*$, with high effective curvature. Investment $I(t)$ is positive and growing. The distributed mode is non-viable: $T_{dist} > T^*$. The trajectory moves rightward in (ρ, T) space as the technology matures (applications with higher ρ become addressable).*

Phase II: Frenzy *($t \in [t_1, t_2]$). Speculative financing raises the effective information temperature of the financial system supporting centralized production. The trajectory moves upward toward T^*. Overinvestment per Theorem 4.1 accelerates cumulative production.*

Phase III: Turning point *($t = t_2$). The system undergoes a* saddle-node bifurcation*: the high-T fixed point (speculative financial equilibrium) collides with the unstable equilibrium separating it from the low-T fixed point (fundamental-value equilibrium). At the bifurcation, small perturbations trigger a discontinuous jump from the speculative to the fundamental equilibrium. This is the crisis.*

Formally, the financial system's effective temperature satisfies:

$$\dot{T}_{\text{fin}} = f(T_{\text{fin}}, Q) = \underbrace{\beta \cdot I(Q)}_{\text{speculation}} - \underbrace{\mu \cdot K_{\text{eff}}(T_{\text{fin}})}_{\text{stabilization}} \tag{24}$$

where β captures the rate at which investment generates speculative froth and μ captures the stabilizing effect of curvature. The saddle-node occurs when $f = 0$ and $\partial f / \partial T = 0$ simultaneously.

Phase IV: Deployment *($t \in [t_2, t_3]$). Post-crisis, the financial system operates at low T (fundamental values). Meanwhile, $T_{dist}(Q)$ has fallen below T^* due to cumulative learning during the installation and frenzy phases. Distributed production enters and expands. The trajectory moves leftward in (ρ, T) space as distributed producers address increasingly complementary tasks (lower ρ).*

Phase V: Maturity *($t > t_3$). Both centralized and distributed modes coexist, with centralized production dominant for high-ρ (substitutable, scale-economy) tasks and distributed production dominant for low-ρ (complementary, diversity-premium) tasks. The (ρ, T) trajectory converges to a fixed point determined by the technology's learning rate exhaustion ($\alpha_{\text{eff}} \to 0$).*

Proof sketch. The phase characterization follows from the dynamics of the three state variables $(Q, T_{\text{fin}}, T_{\text{dist}})$.

Installation: Initial conditions are $Q = Q_0$ small, $T_{\text{fin}} = T_{\text{fin},0}$ moderate, $T_{\text{dist}} = T_{\text{dist},0} > T^*$. Positive investment increases Q. Financial temperature T_{fin} remains below T^* because the stabilization term μK_{eff} dominates.

Frenzy: As Q grows, investment $I(Q)$ increases (firms chase the learning curve), and the speculation term βI grows faster than the stabilization term μK_{eff}. This is because K_{eff} is bounded while I is unbounded in the Nash equilibrium. T_{fin} rises.

Turning point: The dynamics of T_{fin} admit two fixed points when βI is moderate (a stable low-T and a stable high-T equilibrium separated by an unstable intermediate). As I increases, the two stable points approach each other. At the saddle-node bifurcation, the high-T equilibrium and the unstable point merge and annihilate, leaving only the low-T equilibrium. The system jumps discontinuously—this is the crisis.

Deployment: Post-bifurcation, T_{fin} is low. Meanwhile, cumulative Q accumulated during Phases I–III has reduced T_{dist} below T^* (Theorem 5.1). Distributed entry occurs.

Maturity: Learning exhaustion (Q large, $c'(Q) \to 0$) stabilizes the cost curve. The bifurcation condition for the next technology begins to emerge. $\square$

Remark 7.2. *The turning point is a* discontinuous *transition, not a gradual one. This is why technology transitions pass through crises rather than smooth adjustments. The saddle-node bifurcation is structurally stable—small perturbations to the model parameters change the timing but not the existence or character of the discontinuity. This explains why crises are a robust feature of technology cycles despite varying institutional contexts.*

8 Duration and Amplitude

Proposition 8.1 (Cycle duration). *The duration of the full technology cycle (installation through maturity) scales as:*

$$\tau \sim \frac{1}{\alpha} \ln\left(\frac{c_0}{c^*}\right) \cdot \frac{1}{1 + (N-1)\alpha\phi/(r+\delta)} \tag{25}$$

where $c^* = g^{-1}(T^*/T_0)$ is the unit cost at which distributed production becomes viable.

Proof. The crossing time requires $Q(\hat{t}) = (c_0/c^*)^{1/\alpha}$. With aggregate investment rate $N \cdot I^{\text{Nash}} \approx N \cdot I^{\text{coop}} \cdot [1 + (N-1)\alpha\phi/(r+\delta)]$ and $\dot{Q} \approx N \cdot I^{\text{Nash}}$, integrating gives $\hat{t} \propto Q^*/\dot{Q}$. Taking logs: $\ln Q^* = (1/\alpha) \ln(c_0/c^*)$. $\square$

Corollary 8.2 (Duration compression). *If successive technologies have learning rates $\alpha_1 < \alpha_2 < \cdots$ and comparable cost ratios c_0/c^*, then cycle durations decrease:*

$$\frac{\tau_{n+1}}{\tau_n} \approx \frac{\alpha_n}{\alpha_{n+1}} \cdot \frac{1 + (N_n - 1)\alpha_n\phi_n/(r+\delta)}{1 + (N_{n+1} - 1)\alpha_{n+1}\phi_{n+1}/(r+\delta)} \tag{26}$$

When $\alpha_{n+1} > \alpha_n$ and $N_{n+1} \geq N_n$, both factors are less than 1, producing compression.

Remark 8.3. *This provides a structural explanation for Fact 5: successive technology cycles are shorter not because of some exogenous acceleration of time, but because (i) more information-intensive technologies have higher learning rates α, and (ii) later technologies operate in markets with more competitors N and higher spillover rates ϕ, all of which accelerate the crossing.*

Proposition 8.4 (Cycle amplitude). *The amplitude of overinvestment during the installation-frenzy phase is:*

$$A = K \cdot \frac{N-1}{N} \cdot \frac{\alpha\phi}{r + \delta - \alpha\phi} \tag{27}$$

measuring the present value of excess investment as a fraction of the curvature premium. The amplitude is increasing in K (complementarity of the technology), N (number of competitors), α (learning rate), and ϕ (spillover rate).

Proof. The excess investment per firm is $(N-1)\alpha\phi/(r+\delta)$ times the cooperative rate, per Theorem 4.1. Total excess across N firms over the cycle duration τ is approximately $N \cdot (N-1)\alpha\phi/(r+\delta) \cdot I^{\text{coop}} \cdot \tau$. Normalizing by the curvature premium $K \cdot \Delta_{\text{div}}/r$ and substituting the duration formula yields (27) after simplification. □

9 Historical Calibration

We calibrate the model against five technology transitions, using the learning rate α and the complementarity parameter ρ as the primary technology-specific inputs.

9.1 Parameter Identification

For each technology, we identify:

- α: from engineering learning curve studies (cost as a function of cumulative production).

- ρ: from the production structure (how many heterogeneous inputs are required, and how substitutable they are).

- N: from the market structure during the installation phase.

- Cycle timing: installation start, frenzy peak, crisis, deployment onset, maturity.

9.2 Five Transitions

Table 1: Technology cycle parameters and predictions

Technology	α	ρ	K	N	Duration (years)	Predicted τ	Actual τ
Railroads	0.18	−2.0	0.71	8	1840–1890	52	50
Electrification	0.22	−1.0	0.50	12	1890–1930	42	40
Telephony	0.20	−0.5	0.38	5	1880–1920	44	40
Internet	0.30	0.0	0.25	20	1990–2010	18	20
AI (projected)	0.35	varies	varies	8	2020–?	12–15	—

9.2.1 Railroads (1840–1890)

The railroad is the canonical technology cycle. The learning rate $\alpha \approx 0.18$ is estimated from the decline in construction cost per mile as cumulative mileage increased (Fogel, 1964). Complementarity is strong ($\rho \approx -2$): a railroad requires coordinated investment in track, rolling stock, stations, signaling, and trained personnel—failure of any one component renders the others worthless.

Installation (1840–1860): Railroad construction required concentrated capital on a scale unprecedented for private enterprise, driving the creation of modern capital markets (Baskin, 1988). $N \approx 8$ major railroad systems competed for trunk routes.

Frenzy (1860–1873): Track mileage doubled while freight rates fell. Financial innovation (mortgage bonds, preferred stock) expanded the investor base, raising T_{fin}.

Turning point (1873): The failure of Jay Cooke & Company triggered a banking panic. The crisis sequence matches the theorem: financial crisis (bank failures) preceded production disruption (rate wars among railroads) and governance restructuring (Interstate Commerce Act, 1887).

Deployment (1880–1910): With infrastructure in place and construction costs dramatically reduced, small shippers and manufacturers gained access to national markets. The distributed adoption phase generated the Second Industrial Revolution.

The model predicts duration $\tau \approx 52$ years using (25) with $c_0/c^* \approx 8$ (cost per ton-mile ratio), compared to the actual approximately 50 years.

9.2.2 Electrification (1890–1930)

Learning rate $\alpha \approx 0.22$ from the decline in electricity generation cost per kWh (Fouquet, 2014). Complementarity $\rho \approx -1$: electrification requires coordinated investment in generation, transmission, distribution, and end-use equipment, but components are somewhat more modular than railroads.

Installation: Required regulated utility monopolies with guaranteed returns to justify generation and transmission investment. $N \approx 12$ regional utility systems.

Frenzy: Utility holding companies (Insull empire, Associated Gas and Electric) used financial leverage to expand. T_{fin} rose as leverage ratios exceeded 10:1.

Turning point (1929–32): Utility holding company collapses followed the stock market crash. Crisis sequence: stock crash (financial), industrial production collapse (productive), regulatory restructuring under SEC and PUHCA (governance).

Deployment: Rural electrification (REA, 1935 onward) brought electricity to distributed users at dramatically lower cost, enabling the appliance revolution and suburban manufacturing.

Model prediction: $\tau \approx 42$ years; actual approximately 40 years.

9.2.3 Telephony (1880–1920)

Learning rate $\alpha \approx 0.20$ from the decline in per-line installation and switching cost (Mueller, 1997). The complementarity $\rho \approx -0.5$ reflects moderate input heterogeneity: switching equipment, copper lines, operator training, and billing systems must coordinate, but are more modular than heavy infrastructure.

Installation: AT&T's Bell System required monopoly protection (natural monopoly argument) to finance the long-distance network. $N \approx 5$ (Bell plus independents).

Frenzy: Duplication of local networks by competing independents in the 1890s–1900s.

Turning point: Kingsbury Commitment (1913) and subsequent regulation replaced market crisis with regulatory restructuring.

Deployment: Universal service policy extended telephone access from businesses to households. By the 1920s, distributed adoption was self-sustaining.

Model prediction: $\tau \approx 44$ years; actual approximately 40 years.

9.2.4 Internet (1990–2010)

Learning rate $\alpha \approx 0.30$ from the decline in bandwidth cost per Mbps and storage cost per GB (Nagy et al., 2013). Complementarity $\rho \approx 0$ (approximately Cobb-Douglas): the internet combines heterogeneous components (fiber, routers, servers, software, content) but with significant modularity through layered protocols.

Installation (1993–1998): Building the internet backbone required concentrated investment by a small number of telecom carriers and ISPs. $N \approx 20$ in the broader internet ecosystem.

Frenzy (1998–2000): Venture capital and IPO markets financed speculative expansion. T_{fin} rose dramatically as "eyeball" valuations replaced revenue multiples.

Turning point (2000–2002): NASDAQ crash (March 2000), followed by telecom production collapse (WorldCom, Global Crossing), followed by governance restructuring (Sarbanes-Oxley).

Deployment (2003–present): Cloud computing, mobile broadband, and SaaS enabled distributed adoption at near-zero marginal cost. The infrastructure built during the bubble (fiber, data centers) supported decades of growth.

Model prediction: $\tau \approx 18$ years; actual approximately 20 years.

9.2.5 AI (2020–projected)

Learning rate $\alpha \approx 0.35$ from the decline in training cost per parameter (Epoch, 2024). The complementarity parameter ρ *varies by task*: training is near-Leontief ($\rho \to -\infty$) while inference approaches linear substitution ($\rho \to 1$). This is the training-inference bifurcation derived in Smirl (2026d).

Installation (2020–2025): Concentrated investment by $N \approx 8$ hyperscale firms (Microsoft, Google, Amazon, Meta, Oracle, xAI, Anthropic, and others). Training runs exceeding \$100M. Overinvestment ratio ≈ 3–$4\times$ (consistent with Theorem 4.1 for $N = 8$, $\alpha = 0.35$, $\phi \approx 0.5$).

Frenzy (2024–2026?): Speculative AI valuations, capital expenditure growing at $4\times$/year against $1.5\times$ revenue growth. T_{fin} rising.

Turning point (projected): The model predicts a financial adjustment when the correlation robustness of AI-linked portfolios degrades (Theorem 6.1). Duration projection: $\tau \approx 12$–15 years for the full cycle.

Deployment: Distributed inference using models trained during the installation phase. The training-inference bifurcation (Theorem 10.1 below) predicts that training remains cen-

tralized while inference distributes—the first technology cycle where the split is structural rather than just sequential.

10 The Bifurcation Theorem

The AI transition reveals a feature that earlier technology cycles exhibited only weakly: the technology contains tasks with fundamentally different ρ values, leading to structural rather than temporal bifurcation.

Theorem 10.1 (Task bifurcation)**.** *Consider a technology with a spectrum of tasks indexed by $\rho \in [\rho_{\min}, \rho_{\max}]$, where each task has effective curvature $K(\rho) = (1 - \rho)(J - 1)/J$ and the distributed mode has information temperature $T_{dist}(Q)$ decreasing in cumulative investment. At any cumulative production level Q:*

(a) *Tasks with $\rho > \rho^*(Q)$ are distributed, where $\rho^*(Q)$ solves:*

$$K(\rho^*) = \frac{T_{dist}(Q)}{T^*(\rho^*)/K(\rho^*)} = T_{dist}(Q) \cdot \frac{K(\rho^*)}{T^*(\rho^*)} \tag{28}$$

Equivalently, $T_{dist}(Q) = T^*(\rho^*)$: the task-specific crossing condition.

(b) *Tasks with $\rho < \rho^*(Q)$ remain centralized.*

(c) *As Q increases, $\rho^*(Q)$ decreases monotonically: increasingly complementary tasks become distributable.*

(d) *In the limit $Q \to \infty$, $\rho^* \to -\infty$ and all tasks distribute unless T_{dist} has a positive lower bound $\bar{T} > 0$, in which case tasks with $T^*(\rho) < \bar{T}$ remain permanently centralized.*

Proof. Part (a): For task ρ, distributed production is viable when $T_{\text{dist}}(Q) < T^*(\rho)$. Since T^* is increasing in the strength of complementarity (lower ρ means the curvature-to-temperature margin is more generous), tasks with weaker complementarity (higher ρ) cross first.

Part (b): Tasks where $T_{\text{dist}}(Q) > T^*(\rho)$ have $K_{\text{eff}}^{\text{dist}} = 0$, so distributed production achieves no curvature premium. Centralized production with $T_{\text{cent}} < T^*$ retains a positive premium.

Part (c): Monotonicity follows from $\partial T_{\text{dist}}/\partial Q < 0$ (Theorem 5.1).

Part (d): If $T_{\text{dist}}(Q) \to 0$ as $Q \to \infty$, then every finite $T^*(\rho)$ is eventually exceeded. If $T_{\text{dist}} \to \bar{T} > 0$, then tasks with $T^*(\rho) < \bar{T}$ are permanently above the crossing threshold. $\square$

Remark 10.2 (Application to AI)**.** *For AI, the task spectrum spans:*

- Training *($\rho \to -\infty$): Near-Leontief complementarity. Data, compute, and algorithms must synchronize perfectly. K is maximal but T^* is also maximal, so the crossing condition $T_{dist} < T^*$ involves extremely high distributed coordination requirements. Training will be among the last tasks to distribute, if ever.*

- Complex reasoning *($\rho \approx -1$): Strong complementarity between knowledge domains. Distributes once coordination protocols (e.g., mixture-of-experts routing) reduce T_{dist} sufficiently.*

- Routine inference *(ρ ≈ 0): Cobb-Douglas. Distributes early because K is moderate and T^* is moderate.*

- Simple generation *(ρ → 1): Near-linear substitution. Cost determines the winner. Distributes as soon as distributed compute is cheaper per token.*

The prediction is a gradient of decentralization*: simple tasks distribute first, complex tasks last, with training potentially remaining permanently centralized.*

11 Predictions for the AI Transition

The general theory, calibrated to the AI transition's specific parameters ($\alpha \approx 0.35$, $N \approx 8$, $\phi \approx 0.5$, ρ varying by task), generates several testable predictions.

Proposition 11.1 (AI cycle predictions). *Under the model parameters for the AI transition:*

(a) **Overinvestment ratio:** *The Nash equilibrium investment exceeds the social optimum by a factor of approximately 3–4×. Current hyperscaler capital expenditure growth ($\sim 4\times$/year vs. $\sim 1.5\times$ revenue growth) is consistent with this prediction.*

(b) **Financial crisis timing:** *The correlation robustness of AI-linked equity portfolios will degrade before the production value of AI complementarities. Specifically, the diversification benefits of holding multiple AI firms' stock will diminish as their return correlations increase toward 1, preceding any reduction in the productive value of AI applications.*

(c) **Inference cost crossing:** *Distributed inference becomes cost-competitive when:*

$$c_{dist}(Q) \leq c_{cent}(Q) + K_{eff}^{cent} \cdot \Delta_{div} \tag{29}$$

At current learning rates, this crossing occurs approximately 2028–2030 for routine inference tasks ($\rho \approx 0$) and later for complex reasoning ($\rho < 0$).

(d) **Training remains centralized:** *For training ($\rho \to -\infty$), the critical temperature T^* is so high that T_{dist} cannot realistically fall below it unless a fundamentally new coordination technology emerges. Training centralization is structural, not cyclical.*

(e) **Settlement infrastructure demand:** *When distributed inference exceeds a critical mass threshold, the resulting volume of micro-transactions generates demand for settlement infrastructure (stablecoins, real-time payments) that feeds back to the monetary system via the mechanism in Smirl (2026e).*

(f) **Cycle duration:** *The full AI cycle (installation through deployment equilibrium) will last approximately 12–15 years (2020–2032/35), making it the shortest major technology cycle in history—consistent with Corollary 8.2 and the highest observed α.*

12 Relation to Existing Literature

Schumpeter (1942) and creative destruction. Schumpeter's insight that capitalism generates endogenous instability through innovation is the starting point. The present paper micro-founds the creative destruction mechanism: it is not merely that new technologies replace old ones, but that the *investment required to develop* the new technology undermines the *organizational structure* that financed it. The CES curvature parameter ρ determines *what kind* of destruction occurs (productive complementarities vs. organizational forms), while the learning rate α determines *how fast*.

Perez (2002) and techno-economic paradigms. Perez's five-phase model is the most detailed empirical description of technology cycles. The present paper derives these phases rather than positing them. The key advance is identifying the turning point as a saddle-node bifurcation (Theorem 7.1), which explains both its discontinuous character and its structural stability across different institutional contexts.

Kondratiev (1925) and long waves. Kondratiev's 50–60 year periodicity is an empirical observation. The present framework explains both the periodicity (via the duration formula (25)) and its compression over time (via increasing α). Importantly, the theory does not predict a fixed period—the period depends on α, N, ϕ, and c_0/c^*, all of which are technology-specific.

Arrow (1962) and learning by doing. Arrow's foundational insight that unit costs decline with cumulative production is one of our three primitives. The present paper adds the interaction with CES complementarity and information frictions: learning not only reduces cost but also shifts the viability boundary for distributed alternatives by changing the effective information temperature.

Aghion and Howitt (1992) and endogenous growth. The Schumpeterian growth literature models innovation as replacing incumbents. The present paper focuses on a different margin: not replacement of one technology by another, but the *organizational transition* from centralized to distributed production of the *same* technology. This is precisely the Perez deployment phase, which existing endogenous growth models typically abstract away.

Acemoglu and Restrepo (2018) and automation. The AI-specific predictions connect to the automation literature through the task bifurcation theorem (Theorem 10.1): tasks with high ρ (substitutable, routine) distribute and potentially automate first, while tasks with low ρ (complementary, requiring diverse coordination) resist both centralization and automation.

Wright's Law and learning curves. The empirical learning curve literature (Wright, 1936; Nagy et al., 2013; Farmer and Lafond, 2016) provides the α estimates that drive the

quantitative calibration. The present paper does not contribute new learning curve estimates but shows how the learning rate interacts with CES complementarity and information frictions to produce cycle dynamics.

Companion papers. The present paper builds directly on three companions:

- Smirl (2026a): Provides the CES triple role theorem (superadditivity, correlation robustness, strategic independence all governed by K).

- Smirl (2026b): Provides the free energy framework ($\mathcal{F} = \Phi - TH$) that parameterizes information frictions.

- Smirl (2026c): Provides the effective curvature theorem ($K_{\text{eff}} = K(1 - T/T^*)^+$) and the non-uniform degradation result that drives the crisis sequence.

13 Conclusion

This paper has shown that the technology cycle—the recurring pattern of concentrated investment, overinvestment, crisis, and distributed deployment observed across two centuries of technological change—is not a contingent historical regularity but a mathematical consequence of three well-documented primitives: CES complementarity, learning-by-doing, and information frictions.

The central mechanism is self-undermining: concentrated investment finances the learning curve that reduces the information temperature required for distributed alternatives, so the organizational form that funds the technology creates the conditions for its own displacement. This is not a claim that centralized structures are inefficient—on the contrary, they are *necessary* during the installation phase when $T_{\text{dist}} > T^*$ and only centralized coordination can exploit the technology's curvature. The claim is that their very success undermines the conditions for their dominance.

Three results deserve emphasis. First, the crisis is a bifurcation, not a gradual adjustment. The saddle-node structure of the turning point (Theorem 7.1) means that the transition from speculative to fundamental equilibrium is necessarily discontinuous—there is no smooth path from installation to deployment. This explains why technology transitions always pass through crises, despite varying institutional contexts and policy responses.

Second, the crisis has a universal sequence. Correlation robustness fails first (financial crisis), followed by superadditivity (production disruption), followed by strategic independence (governance restructuring). This sequence, derived from the quadratic-vs-linear degradation rates in the effective curvature theorem, matches the observed order in all five historical transitions examined.

Third, successive cycles compress because learning rates increase. As technologies become more information-intensive, α rises, and the duration formula $\tau \sim (1/\alpha) \ln(c_0/c^*)$ shrinks. The AI transition, with $\alpha \approx 0.35$—the highest observed for any major technology—should produce the shortest cycle yet, completing in approximately 12–15 years.

For the AI transition specifically, the model predicts a structural bifurcation that earlier cycles lacked: training remains centralized (near-Leontief complementarity, extremely high

T^*) while inference progressively distributes (moderate complementarity, declining T_{dist}). This is not a temporary phase but a permanent structural feature of the technology, distinguishing the AI cycle from its predecessors where centralized and distributed modes competed over the same tasks.

The framework is deliberately parsimonious—three primitives, two technology-specific parameters (α and ρ)—and generates quantitative predictions that can be tested against both historical data and the unfolding AI transition. The overinvestment ratio, crisis sequence, duration compression, and task bifurcation are all observable. If the AI transition follows the predicted pattern, with financial stress in AI-linked assets preceding productive adjustment and governance restructuring, it would provide real-time confirmation of a theory calibrated against 200 years of prior transitions.

References

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

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60(2): 323–351.

Arrow, Kenneth J. 1962. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29(3): 155–173.

Baskin, Jonathan Barron. 1988. "The Development of Corporate Financial Markets in Britain and the United States, 1600–1914: Overcoming Asymmetric Information." *Business History Review* 62(2): 199–237.

Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. "General Purpose Technologies: 'Engines of Growth'?" *Journal of Econometrics* 65(1): 83–108.

David, Paul A. 1990. "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox." *American Economic Review Papers and Proceedings* 80(2): 355–361.

Epoch AI. 2024. "Trends in Machine Learning." Technical Report. https://epochai.org/trends.

Farmer, J. Doyne, and François Lafond. 2016. "How Predictable Is Technological Progress?" *Research Policy* 45(3): 647–665.

Fogel, Robert William. 1964. *Railroads and American Economic Growth: Essays in Econometric History.* Baltimore: Johns Hopkins Press.

Fouquet, Roger. 2014. "Long-Run Demand for Energy Services: Income and Price Elasticities over Two Hundred Years." *Review of Environmental Economics and Policy* 8(2): 186–207.

Kondratiev, Nikolai D. 1925. "The Major Economic Cycles." Translated in *Review of Economic Statistics* 17(6): 105–115 (1935).

Mueller, Milton L. 1997. *Universal Service: Competition, Interconnection, and Monopoly in the Making of the American Telephone System*. Cambridge, MA: MIT Press.

Nagy, Béla, J. Doyne Farmer, Quan M. Bui, and Jessika E. Trancik. 2013. "Statistical Basis for Predicting Technological Progress." *PLoS ONE* 8(2): e52669.

Perez, Carlota. 2002. *Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages*. Cheltenham, UK: Edward Elgar.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper & Brothers.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50(3): 665–690.

Smirl, Jon. 2026a. "The CES Triple Role: Superadditivity, Correlation Robustness, and Strategic Independence from a Single Curvature Parameter." Working Paper.

Smirl, Jon. 2026b. "Free Energy Economics: A Unified Framework from CES Aggregation and Shannon Entropy." Working Paper.

Smirl, Jon. 2026c. "Production Under Information Frictions: A CES Free Energy Theory of the Firm." Working Paper.

Smirl, Jon. 2026d. "Endogenous Decentralization: A Differential Game of Learning Curves, AI Hardware Crossing, and Self-Sustaining Distributed Adoption." Working Paper.

Smirl, Jon. 2026e. "The Settlement Feedback: Stablecoin Demand, Treasury Markets, and Monetary Policy in the Mesh Era." Working Paper.

Ward, J. R. 1974. *The Finance of Canal Building in Eighteenth-Century England*. Oxford: Oxford University Press.

Wright, Theodore P. 1936. "Factors Affecting the Cost of Airplanes." *Journal of the Aeronautical Sciences* 3(4): 122–128.