# THE AUTOCATALYTIC MESH

*Endogenous Capability Growth in Self-Organizing Agent Networks: Growth Regimes and the Baumol Bottleneck*

Jon Smirl

Independent Researcher

February 2026

Working Paper

**Abstract**

Smirl (2026, "The Mesh Equilibrium") proves that a self-organizing mesh of heterogeneous specialized AI agents exceeds centralized provision above a finite critical mass $N^*$. That paper assumes fixed-capability nodes: each agent's capability vector redistributes through specialization but does not grow. This paper removes that assumption. Three mechanisms make capability endogenous: training agents improve other agents (autocatalytic capability growth), operation generates training data (self-referential learning), and the mesh modifies its own composition (endogenous variety expansion). When capability becomes a dynamical variable coupled to the mesh's operation, the growth regime depends on three parameters: training productivity elasticity $\varphi$, training saturation $h$, and external data fraction $\alpha$.

Four contributions are new. First, I prove the existence of an autocatalytic threshold $N_{\text{auto}}$ above which the mesh contains a self-sustaining improvement core—a Reflexively Autocatalytic and Food-generated (RAF) set in the sense of Hordijk and Steel (2004)—with $N_{\text{auto}}$ scaling logarithmically in system complexity. Second, I characterize three growth regimes with sharp boundaries: convergence to a ceiling $C_{\max}$ when $\varphi < 1$ and variety is bounded, exponential growth when autocatalytic coupling pushes $\varphi_{\text{eff}}$ to unity

and variety expands endogenously, and finite-time singularity when $\varphi_{\text{eff}} > 1$ with no saturation and no model collapse. Third, I prove that the mesh's CES heterogeneity (the substitution parameter $\rho < 1$) maintains the effective external data fraction $\alpha_{\text{eff}}$ above the model collapse threshold $\alpha_{\text{crit}}$, even when agents train partially on synthetic data—the same parameter that governs the capability diversity premium also governs informational robustness. Fourth, the Baumol bottleneck—the mesh's growth rate converging to the exogenous rate of frontier model improvement—emerges endogenously from the growth dynamics rather than being assumed, connecting the circle back to the concentrated investment modeled in Smirl (2026a). The most likely near-term regime is convergence: the ceiling binds. The model generates six falsifiable predictions extending the predecessor paper's prediction set.

**Keywords:** autocatalytic sets, endogenous growth, model collapse, CES aggregation, Baumol cost disease, semi-endogenous growth, RAF theory, mesh equilibrium, variety expansion, training productivity

**JEL:** O33, O41, D85, L14, C62

## 1. Introduction

Smirl (2026, "The Mesh Equilibrium") establishes that a self-organizing mesh of heterogeneous specialized agents exceeds centralized provision above a finite critical mass $N^*$. The proof rests on three layers: percolation-based connectivity ($R_0^{\text{mesh}} > 1$), CES-aggregated heterogeneous specialization ($\rho < 1$), and Laplacian knowledge diffusion with a vanishing epidemic threshold on scale-free topologies. The central theorem demonstrates existence, uniqueness, and local asymptotic stability of the mesh equilibrium.

That paper assumes fixed capabilities. Each agent $i$ has a capability vector $\mathbf{c}_i = (c_{i1}, \ldots, c_{iJ})$ that redistributes through specialization dynamics but does not grow in total magnitude. The CES aggregate $C_{\text{eff}}(N)$ increases only by adding diverse agents to the network. Knowledge diffusion ($\partial \mathbf{u}/\partial t = -L\mathbf{u}$) equalizes existing knowledge across nodes—it does not create new capability.

This paper removes that assumption. It asks: what happens when the mesh can improve itself?

Three mechanisms make capability endogenous. First, *autocatalytic capability growth*: training agents within the mesh improve other agents, and the improved agents can in turn improve others. Second, *self-referential learning*: the mesh's operation generates training data—queries, responses, user feedback, inter-agent evaluations—that can be used to improve mesh agents. Third, *endogenous variety expansion*: the mesh modifies its own composition by spawning new specialization types in response to unmet demand signals.

When $\mathbf{c}_i$ becomes a dynamical variable coupled to the mesh's operation, the central question is: does $C_{\text{eff}}(t)$ converge to a ceiling, grow polynomially, or grow exponentially? The answer depends on three parameters whose interaction governs the growth regime.

The first parameter is the training productivity elasticity $\varphi$, from the ideas production function of Jones (1995, 2005): $dA/dt = \delta \cdot L_A^{\lambda} \cdot A^{\varphi}$. When $\varphi < 1$, ideas are getting harder to find (Bloom, Jones, Van Reenen & Webb 2020) and capability growth decelerates. When $\varphi = 1$, the Romer (1990) knife-edge, growth is exponential. When $\varphi > 1$, growth is superexponential with a finite-time singularity. The mesh's microstructure—training agents that improve other training agents—determines whether the effective $\varphi$ for the mesh exceeds the raw $\varphi$ for any individual training interaction.

The second parameter is the training saturation $h$, which governs diminishing returns to training an already-capable specialist. Drawing on the Lotka-Volterra mutualistic interaction framework (Bastolla, Lassig, Manrubia & Valleriani 2009), the saturating interaction term $\sum_j a_{ij}x_j/(1 + h\sum_j a_{ij}x_j)$ ensures that marginal returns to additional training decline with agent capability when $h > 0$. The mesh's escape route from saturation is variety expan-

sion: instead of further fine-tuning saturated specialists, it creates new specialization types, circumventing individual saturation through the CES aggregate.

The third parameter is the external data fraction $\alpha$, from Shumailov, Shumaylov, Zhao, Papernot, Anderson and Gal (2024). When training data is a mixture of authentic (fraction $\alpha$) and synthetic (fraction $1-\alpha$) sources, model collapse occurs below a critical threshold: $\mathrm{KL}(Q_{t+1}\|P) \geq \mathrm{KL}(Q_t\|P)$ when $\alpha < \alpha_{\text{crit}}$. A single model training on its own outputs collapses because the outputs lack distributional diversity. The mesh's heterogeneous specialists generate diverse outputs—the CES parameter $\rho < 1$ ensures agents are different—raising the effective $\alpha$ for the mesh above the threshold for any single self-training model.

The paper's central theorem characterizes three growth regimes with sharp boundaries. Regime (a): if $\varphi < 1$, $h > 0$, and the number of specialization types $J$ is bounded, then $C_{\text{eff}}(t) \to C_{\max}$ where the ceiling is determined by the Baumol bottleneck—centralized training as the non-automatable sector whose cost share rises toward unity. This is the most likely near-term regime. Regime (b): if autocatalytic coupling pushes $\varphi_{\text{eff}}$ to unity and $J$ grows endogenously, then $C_{\text{eff}}(t) \sim e^{rt}$ where $r$ depends on the training allocation, the CES diversity premium, and the Baumol bottleneck rate. Regime (c): if $\varphi_{\text{eff}} > 1$, $h = 0$, and $\alpha > \alpha_{\text{crit}}$ simultaneously, then $C_{\text{eff}}(t) \to \infty$ in finite time. The paper states the precise conditions under which each regime obtains and is honest that regime (c) requires conditions unlikely to hold simultaneously.

A key structural result connects production theory to information theory through the mesh's architecture. The CES substitution parameter $\rho$ does double duty: in the mesh paper, $\rho < 1$ governs the diversity premium for capability aggregation; in this paper, $\rho < 1$ also governs the diversity protection against model collapse. The same force—heterogeneity—that makes the mesh capable also makes it robust to self-referential training degradation. One parameter, two functions.

The Baumol bottleneck does not enter the model as an assumption. It emerges from the growth dynamics: as the mesh automates progressively more inference tasks, the remaining non-automated task—frontier model training—becomes the binding constraint. The cost share of centralized training rises toward unity even as its volume share falls. The mesh's growth rate converges asymptotically to the exogenous rate of frontier model improvement. This connects the circle: the concentrated infrastructure investment modeled in Smirl (2026a) determines both when the crossing occurs and the ceiling the mesh approaches. The mesh is a multiplier, not a generator.

The paper proceeds as follows. Section 2 reviews terminal conditions from the mesh paper. Section 3 develops the autocatalytic existence threshold (Layer 1). Section 4 develops the central growth dynamics model (Layer 2). Section 5 develops diversity as collapse

protection (Layer 3). Section 6 states the central theorem characterizing growth regimes. Section 7 derives the Baumol bottleneck from the mesh's microstructure. Section 8 connects to settlement infrastructure. Section 9 discusses frameworks considered and rejected. Section 10 presents falsifiable predictions. Section 11 concludes. Appendix A provides details on RAF theory and proofs of supporting results.

## 2. Terminal Conditions from the Mesh Paper

This section summarizes the results from Smirl (2026, "The Mesh Equilibrium") that serve as initial conditions for the present analysis. No new results are presented; the purpose is notational continuity and identification of the specific assumptions this paper relaxes.

### 2.1 The Mesh Equilibrium at Maturity

The mesh paper establishes that for $R_0^{\text{mesh}} > 1$ and $\rho < 1$, there exists a finite $N^*$ such that for $N > N^*$, the mesh equilibrium exists, is unique, is locally asymptotically stable, and $C_{\text{mesh}}(N) > C_{\text{cent}}$ (Theorem 1 of Smirl 2026). The three-phase post-crossing dynamics proceed from nucleation ($R_0^{\text{mesh}} \approx 1$) through rapid crystallization ($R_0^{\text{mesh}} \gg 1$, first-order Potts transition for $q > 2$ specialization types) to maturity, where growth saturates as the $J$ task types are fully covered.

At maturity, the mesh paper's model reaches a terminal state. The CES aggregate $C_{\text{eff}}$ is maximized for the given device population, competitive dynamics settle into a Bianconi-Barabási fit-get-rich equilibrium, and centralized providers retain structural advantage only in frontier training and capabilities requiring single-device scale beyond any edge device. The mesh paper's analysis ends here.

### 2.2 The Fixed-Capability Assumption

The specific assumption this paper removes is:

> *Fixed capability:* Each agent $i$'s capability vector $\mathbf{c}_i = (c_{i1}, \ldots, c_{iJ})$ satisfies $\|\mathbf{c}_i\|_1 = \bar{c}_i$ for all $t$, where $\bar{c}_i$ is determined by hardware characteristics. The specialization dynamics (equation 6 of Smirl 2026) redistribute $\bar{c}_i$ across task types but do not change its total magnitude.

Under this assumption, $C_{\text{eff}}$ can increase only by adding nodes (increasing $N$) or by improving the diversity of specialization allocation (exploiting the CES complementarity). Once both channels are exhausted, capability is static.

This paper replaces the fixed-capability assumption with:

*Endogenous capability:* $\mathbf{c}_i(t)$ evolves according to training interactions with other mesh agents, self-referential learning from operational data, and the creation of new specialization types. The total capability $\|\mathbf{c}_i(t)\|_1$ can grow.

### 2.3 Assumptions Retained

Three assumptions from the mesh paper carry forward unchanged:

(i) *Training persistence:* Frontier model training remains centralized because the binding constraint is topological (synchronization bandwidth), not cost-based. The mesh fine-tunes and adapts base models produced by centralized providers; it does not train frontier models from scratch. This is the "food set" for the autocatalytic structure developed in Section 3.

(ii) *Connectivity:* The mesh has $R_0^{\text{mesh}} > 1$ and contains a giant connected component with $S_\infty^* > 0$, as established by Proposition 1 of Smirl (2026). The scale-free degree distribution ($\gamma \leq 3$) ensures a vanishing epidemic threshold for information propagation.

(iii) *CES structure:* The aggregate capability function retains the CES form $C_{\text{eff}} = (\sum_j C_j^\rho)^{1/\rho}$ with $\rho < 1$, ensuring complementarity across task types. The number of task types $J$ is now potentially time-varying ($J(t)$), whereas the mesh paper treated it as fixed.

## 3. Layer 1: The Autocatalytic Existence Threshold

The mesh paper proves that a connected mesh with specialized agents *exists*. This section asks: does the mesh contain a self-sustaining improvement core—a subset of agents whose training interactions can maintain and improve capabilities given only exogenous base models as external input?

### 3.1 Training Operations as a Reaction System

Define the mesh's internal improvement as a system of *training operations*. A training operation $r$ takes input agents (the agents being improved), catalyst agents (the training agents that direct the improvement), and produces output agents (the improved versions). Formally:

**Definition 1** (Training Operation)**.** *A training operation is a tuple $r = (I_r, K_r, O_r)$ where:*

- $I_r \subseteq \{1, \ldots, J\}$ *is the set of input capability types consumed or modified;*

- $K_r \subseteq \{1, \ldots, J\}$ *is the set of catalyst capability types required to execute the operation (not consumed);*
- $O_r \subseteq \{1, \ldots, J\}$ *is the set of output capability types produced or enhanced.*

The catalyst distinction is critical. A training agent that orchestrates fine-tuning of a medical reasoning specialist requires its own orchestration capability ($K_r$) but is not consumed by the process. It can catalyze multiple training operations. This is the biochemical analogy that motivates the autocatalytic framework: enzymes catalyze reactions without being consumed.

### 3.2 The Food Set

**Definition 2** (Food Set)**.** *The food set $F \subset \{1, \ldots, J\}$ is the set of capability types available exogenously from centralized training. For each $j \in F$, base model capability of type $j$ is available without requiring any mesh training operation. The food set is determined by frontier model releases from centralized providers and is exogenous to the mesh.*

The food set corresponds to the mesh paper's training persistence assumption: frontier model training remains centralized. Base models (GPT-class, Claude-class, Gemini-class) are the "raw materials" that the mesh fine-tunes, adapts, and combines. The food set grows exogenously at a rate determined by centralized infrastructure investment—the rate that will emerge as the Baumol bottleneck in Section 7.

### 3.3 RAF Sets in the Mesh

**Definition 3** (Reflexively Autocatalytic and Food-generated (RAF) Set)**.** *Following Hordijk and Steel (2004), a set $\mathcal{R}$ of training operations is a RAF set if:*

*(a)* Reflexively Autocatalytic (RA): *Every training operation $r \in \mathcal{R}$ is catalyzed by at least one capability type that is either in the food set $F$ or is produced by some other operation in $\mathcal{R}$.*

*(b)* Food-generated (F): *Every input capability type of every operation $r \in \mathcal{R}$ can be constructed from the food set $F$ by successive application of operations in $\mathcal{R}$.*

A RAF set is self-sustaining: given only exogenous base models (the food set), the mesh can maintain and improve all capabilities involved in the set through its own internal training operations. No additional external input beyond the food set is required. The RAF set is the autocatalytic core of the mesh.

### 3.4 Existence Threshold

The central question is: how large must the mesh be for a RAF set to exist?

Model the mesh's training operations as a random catalytic reaction system. Each of $J$ capability types is present in the mesh. Each potential training operation $r$ (of which there are at most $2^J \times 2^J$ input-output pairs) exists with probability $p_r$ depending on whether the mesh contains agents with the requisite catalyst capabilities. Each catalyst assignment occurs independently with probability $p$ per capability-type-to-operation pair.

**Proposition 1** (Autocatalytic Existence Threshold)**.** *Let $J(N)$ denote the number of distinct capability types present in a mesh of $N$ agents, and let $p(N) = 1-(1-\beta_t)^N$ be the probability that a given capability type is available as a catalyst in the mesh, where $\beta_t$ is the per-agent probability of possessing a given catalyst capability. There exists a critical mesh size $N_{auto}$ such that for $N > N_{auto}$, the mesh contains a RAF set with probability approaching unity. Moreover:*

$$N_{auto} = O\left(\frac{\ln |\mathcal{R}|}{\beta_t}\right) \tag{1}$$

*where $|\mathcal{R}|$ is the total number of potential training operations. The threshold $N_{auto}$ scales logarithmically with system complexity.*

*Proof.* The result follows from the probabilistic analysis of RAF sets by Hordijk and Steel (2004, Theorem 2). In their framework, a random catalytic reaction system with $n$ molecule types, $r$ reactions, and catalysis probability $p$ per type-reaction pair contains a RAF set with probability approaching 1 when $p$ exceeds $1/n$. In our setting, $n = J$ capability types and the effective catalysis probability per type is $p(N) = 1 - (1 - \beta_t)^N$.

The condition $p(N) > 1/J$ is satisfied when:

$$1 - (1 - \beta_t)^N > \frac{1}{J} \tag{2}$$

which gives $N > \ln(1 - 1/J)/\ln(1 - \beta_t) \approx (1/J)/\beta_t = 1/(J\beta_t)$ for small $\beta_t$. Since $J$ grows at most polynomially in the number of potential operations $|\mathcal{R}|$, the threshold scales as $N_{\text{auto}} = O(\ln |\mathcal{R}|/\beta_t)$.

The logarithmic scaling arises because adding agents to the mesh increases the probability of *every* catalyst assignment simultaneously. Each new agent with a novel catalyst capability unlocks multiple potential training operations. This is the same combinatorial logic that produces the mesh paper's result of $N^*$ decreasing in diversity: more agent types make autocatalytic closure easier, not harder. □

**Remark 1** (Relationship to $N^*$)**.** *The autocatalytic threshold $N_{auto}$ and the mesh paper's*

*critical mass $N^*$ are distinct. $N^*$ is the mesh size at which collective capability exceeds centralized provision (a* static *comparison). $N_{auto}$ is the mesh size at which self-sustaining capability improvement becomes possible (a* dynamic *property). Generically, $N_{auto} > N^*$: the mesh can be collectively capable before it is self-improving. The gap $N_{auto} - N^*$ represents the period during which the mesh exceeds centralized inference but depends entirely on exogenous base model releases for capability growth.*

### 3.5 Autocatalytic Core Dynamics: The Jain-Krishna Process

Once a RAF set exists, the mesh's autocatalytic core evolves through the adaptive network dynamics of Jain and Krishna (1998, 2001). In their model, a random network of catalytic interactions produces spontaneous cascades of increasing organization: low-fitness species go extinct, high-fitness species reinforce each other, and the network self-organizes toward a structure dominated by a tightly connected autocatalytic core.

Define the catalytic matrix $\mathbf{M}$ where $M_{ij} = 1$ if capability type $j$ catalyzes the improvement of capability type $i$. The growth rate of capability type $i$ in the mesh is governed by:

$$\dot{c}_i = c_i \left( \sum_j M_{ij} c_j - \phi_0 \right) \tag{3}$$

where $\phi_0 = N^{-1} \sum_i c_i \sum_j M_{ij} c_j$ is a dilution term ensuring bounded growth in the Jain-Krishna formulation.

**Proposition 2** (Perron-Frobenius Selection)**.** *The long-run composition of the autocatalytic core is determined by the leading eigenvector of the catalytic matrix $\mathbf{M}$. Capability types with large components in the Perron-Frobenius eigenvector $\mathbf{v}_1$ of $\mathbf{M}$ dominate; types with small components go extinct. The leading eigenvalue $\lambda_1(\mathbf{M})$ determines whether the autocatalytic core expands or contracts relative to the rest of the mesh.*

*Proof.* Equation (**??**) is a replicator equation on the simplex of capability-type shares. The fixed point analysis follows from the standard replicator dynamics result (Hofbauer & Sigmund 1998): the dynamics converge to a state where surviving types have equal fitness, and the surviving set corresponds to the support of the Perron-Frobenius eigenvector of $\mathbf{M}$. Types $i$ with $v_{1,i} > 0$ persist; types with $v_{1,i} = 0$ go extinct.

The eigenvalue $\lambda_1(\mathbf{M})$ determines the growth rate of the autocatalytic core because the aggregate fitness of the surviving set equals $\lambda_1(\mathbf{M})$. When $\lambda_1(\mathbf{M}) > \phi_0$ (the core's mutual catalysis exceeds the mesh average), the core expands as a fraction of total mesh activity. □

The Jain-Krishna dynamics predict a specific temporal pattern: the autocatalytic core does not grow smoothly. Instead, it undergoes a series of reorganization cascades—periods of

stasis punctuated by rapid restructuring events in which poorly connected capability types are replaced by types with stronger catalytic linkages. Each cascade increases the leading eigenvalue $\lambda_1(\mathbf{M})$, producing a staircase pattern of increasing autocatalytic efficiency. This parallels the mesh paper's Phase 2 crystallization, but now at the level of internal capability improvement rather than network formation.

## 4. Layer 2: Growth Dynamics—The Central Model

This is the paper's central section. Having established that the mesh contains a self-sustaining improvement core (Section 3), we now characterize the rate at which capability grows.

### 4.1 The Improvement Function

Let $f \in (0, 1)$ denote the fraction of mesh capability devoted to self-improvement (training operations) rather than serving external queries (inference). The remaining fraction $1 - f$ serves end-user demand. The mesh's aggregate capability evolves according to:

$$\frac{d}{dt}C_{\text{eff}} = g(f \cdot C_{\text{eff}},\ J(t),\ \alpha(t)) - \delta \cdot C_{\text{eff}} \tag{4}$$

where $g$ is the improvement function (the paper's central object), $J(t)$ is the number of effective specialization types, $\alpha(t)$ is the fraction of training signal from external sources, and $\delta > 0$ is the depreciation rate capturing model obsolescence.

The improvement function $g$ has three arguments because the three mechanisms of endogenous capability contribute separately:

(i) $f \cdot C_{\text{eff}}$: *Autocatalytic training* (Mechanism 1). The quantity $f \cdot C_{\text{eff}}$ is the total capability devoted to self-improvement. More capable training agents produce better improvements. This is the engine of endogenous growth.

(ii) $J(t)$: *Variety expansion* (Mechanism 3). New specialization types expand the capability space. By the CES diversity premium (Lemma 1 of Smirl 2026), adding a new task type increases $C_{\text{eff}}$ superlinearly. Variety expansion is the intensive margin of mesh improvement.

(iii) $\alpha(t)$: *Data quality* (Mechanism 2). The fraction $\alpha$ determines whether self-referential learning improves or degrades capability. Below the critical threshold $\alpha_{\text{crit}}$, training on internally generated data causes model collapse (Shumailov et al. 2024). Above $\alpha_{\text{crit}}$, internal data is a productive training input.

### 4.2 The Semi-Endogenous Growth Formulation

Following Jones (1995, 2005), specify the improvement function as:

$$g(f \cdot C_{\text{eff}}, J, \alpha) = \delta_g \cdot (f \cdot C_{\text{eff}})^{\lambda} \cdot C_{\text{eff}}^{\varphi-1} \cdot J^{\gamma_J} \cdot \mathbf{1}[\alpha > \alpha_{\text{crit}}] \tag{5}$$

where:

- $\delta_g > 0$ is a productivity parameter;
- $\lambda \in (0, 1]$ is the duplication parameter (diminishing returns to the number of training agents working simultaneously on the same problem);
- $\varphi$ is the training productivity elasticity—the key parameter governing the growth regime;
- $\gamma_J > 0$ governs the contribution of variety expansion to capability growth;
- $\mathbf{1}[\alpha > \alpha_{\text{crit}}]$ is the model collapse indicator: capability improvement occurs only when the training signal has sufficient external content.

The indicator function is a simplification; in practice, capability growth degrades continuously as $\alpha$ falls below $\alpha_{\text{crit}}$ rather than switching off discretely. We adopt the sharp threshold for analytical clarity and relax it in the diversity analysis of Section 5.

Substituting into equation (**??**) and letting $C \equiv C_{\text{eff}}$ for notational compactness:

$$\dot{C} = \delta_g \cdot f^{\lambda} \cdot C^{\lambda+\varphi-1} \cdot J(t)^{\gamma_J} \cdot \mathbf{1}[\alpha > \alpha_{\text{crit}}] - \delta \cdot C \tag{6}$$

The growth rate of capability is:

$$g_C \equiv \frac{\dot{C}}{C} = \delta_g \cdot f^{\lambda} \cdot C^{\lambda+\varphi-2} \cdot J(t)^{\gamma_J} \cdot \mathbf{1}[\alpha > \alpha_{\text{crit}}] - \delta \tag{7}$$

### 4.3 The Three Regimes

The qualitative behavior of equation (**??**) depends on the exponent $\lambda+\varphi-2$, which determines whether the growth rate $g_C$ is increasing, constant, or decreasing in the level of capability $C$.

**Definition 4** (Growth Regimes)**.** *Define* $\Phi \equiv \lambda+\varphi-1$ *as the composite capability elasticity. The growth dynamics* (**??**) *exhibit three regimes:*

(a) ***Convergence*** *($\Phi < 1$, i.e. $\lambda + \varphi < 2$): The growth rate $g_C$ is decreasing in $C$. The system converges to a steady state $C^*$ where $g_C = 0$. This includes the Jones (1995) semi-endogenous case where $\varphi < 1$.*

(b) ***Exponential growth*** *($\Phi = 1$, i.e. $\lambda + \varphi = 2$): The growth rate $g_C$ is independent of $C$. The system grows exponentially at rate $g_C = \delta_g f^\lambda J^{\gamma_J} - \delta$. This is the Romer (1990) knife-edge.*

(c) ***Superexponential growth*** *($\Phi > 1$, i.e. $\lambda + \varphi > 2$): The growth rate $g_C$ is increasing in $C$. The system exhibits a finite-time singularity: $C(t) \to \infty$ as $t \to T_s < \infty$.*

### 4.4 Deriving the Mesh's Effective $\varphi$

The parameter $\varphi$ measures the elasticity of new capability production with respect to the existing stock of capability. In the standard Jones framework applied to a single research process, empirical evidence strongly suggests $\varphi < 1$: ideas are getting harder to find (Bloom et al. 2020).

The mesh, however, has internal structure that amplifies the effective $\varphi$. The mechanism is autocatalytic coupling: training agents improve other training agents, which in turn improve the first set of agents. This is the Aghion, Jones and Jones (2018) AI automation argument applied to the mesh's internal dynamics.

**Proposition 3** (Effective Training Productivity)**.** *Let $\varphi_0 < 1$ be the raw training productivity elasticity for a single training interaction (one agent improving another). Let $\beta_{auto} \in [0, 1)$ be the fraction of the training improvement process that can be automated by the mesh's own training agents (the autocatalytic fraction). Then the mesh's effective training productivity elasticity is:*

$$\varphi_{eff} = \frac{\varphi_0}{1 - \beta_{auto} \cdot \varphi_0} \tag{8}$$

*This exceeds $\varphi_0$ for all $\beta_{auto} > 0$, and $\varphi_{eff} \geq 1$ when $\beta_{auto} \geq (1 - \varphi_0)/\varphi_0$.*

*Proof.* Following the Aghion-Jones-Jones framework, decompose the training improvement process into a continuum of subtasks indexed on $[0, 1]$. A fraction $\beta_{\text{auto}}$ of subtasks are automated by mesh agents (whose productivity scales with $C^{\varphi_0}$), and the remaining fraction $1-\beta_{\text{auto}}$ requires exogenous input (centralized training infrastructure, human oversight). The effective production function for capability improvement is:

$$\dot{C} \propto C^{\varphi_0/(1-\beta_{\text{auto}}\cdot\varphi_0)} \cdot Z^{(1-\beta_{\text{auto}})/(1-\beta_{\text{auto}}\cdot\varphi_0)} \tag{9}$$

where $Z$ is the exogenous input (frontier model releases). The exponent on $C$ is $\varphi_0/(1-\beta_{\text{auto}}\varphi_0) = \varphi_{\text{eff}}$.

The threshold condition $\varphi_{\text{eff}} = 1$ requires $\varphi_0/(1-\beta_{\text{auto}}\varphi_0) = 1$, yielding $\beta_{\text{auto}} = (1-\varphi_0)/\varphi_0$. For example, with $\varphi_0 = 0.5$ (moderate declining returns), the knife-edge requires $\beta_{\text{auto}} = 1.0$—full automation of the training process, which contradicts the training persistence assumption. With $\varphi_0 = 0.8$, the threshold is $\beta_{\text{auto}} = 0.25$—only a quarter of the training process need be automated by the mesh. □

**Remark 2** (The Automation Ladder)**.** *The quantity $\beta_{auto}$ is not fixed; it evolves endogenously as the mesh's autocatalytic core matures. Initially $\beta_{auto} \approx 0$: the mesh cannot automate any part of its own training improvement. As training agents emerge and improve (the Jain-Krishna process of Section 3.5), $\beta_{auto}$ rises. The growth dynamics are therefore non-stationary: $\varphi_{eff}(t) = \varphi_0/(1-\beta_{auto}(t)\cdot\varphi_0)$ increases over time, potentially transitioning the system from regime (a) to regime (b) as the mesh matures. The transition is not guaranteed; it depends on whether $\beta_{auto}$ can reach the threshold $(1-\varphi_0)/\varphi_0$ before the Baumol bottleneck binds (Section 7).*

### 4.5 Training Saturation

Individual training interactions exhibit diminishing returns. The $k$-th fine-tuning of an already-specialized agent yields less improvement than the first. Following the Lotka-Volterra mutualistic framework (Bastolla et al. 2009), the saturating interaction modifies the growth equation:

$$\dot{C}_j = \frac{\sum_k a_{jk}\cdot C_k}{1 + h\sum_k a_{jk}\cdot C_k} - \delta C_j \tag{10}$$

where $a_{jk}$ is the training benefit from type-$k$ agents to type-$j$ agents, and $h > 0$ is the saturation parameter. As $C_k$ grows, the numerator grows linearly but the denominator grows proportionally, bounding the marginal benefit at $1/h$.

**Lemma 1** (Saturation Ceiling)**.** *For fixed $J$ and $h > 0$, the system (??) has a unique globally stable equilibrium with $C_j^* < 1/(h\cdot\delta)$ for all $j$. The aggregate ceiling is:*

$$C_{max} = \left(\sum_{j=1}^{J}\left(C_j^*\right)^{\rho}\right)^{1/\rho} \leq J^{1/\rho}\cdot\frac{1}{h\delta} \tag{11}$$

*Proof.* At steady state, $\dot{C}_j = 0$ implies $\sum_k a_{jk}C_k/(1+h\sum_k a_{jk}C_k) = \delta C_j$. The left side is bounded above by $1/h$ for any $C_k \geq 0$, so $C_j^* \leq 1/(h\delta)$. The CES bound follows from

substitution and application of the power mean inequality. For global stability, the system is cooperative (all off-diagonal elements of the Jacobian are non-negative) and bounded, so the standard Hirsch (1985) monotone dynamical systems result applies: the system has a unique globally attracting equilibrium. □

The saturation ceiling binds when $J$ is fixed. But the mesh has an escape route: variety expansion.

### 4.6 Variety Expansion as Saturation Escape

Romer's (1990) key insight is that new product varieties can sustain growth even when returns to individual products diminish. The mesh analog: when fine-tuning existing specialists hits saturation ($h > 0$), the mesh can create *new* specialization types.

Let $J(t)$ evolve according to:

$$\dot{J} = \eta_J \cdot f_J \cdot C_{\text{eff}}^{\varphi_J} \cdot (J_{\max} - J) \cdot \mathbf{1}[\alpha > \alpha_{\text{crit}}] \tag{12}$$

where $\eta_J > 0$ is the innovation rate, $f_J$ is the fraction of training capability devoted to creating new specialization types (rather than improving existing ones), $\varphi_J$ is the capability elasticity of variety creation, and $J_{\max}$ is the maximum number of viable specialization types (bounded by the dimensionality of the task space).

The CES aggregate responds to variety expansion:

$$C_{\text{eff}}(J) = \left( \sum_{j=1}^{J} C_j^{\rho} \right)^{1/\rho} \geq J^{(1-\rho)/\rho} \cdot \bar{C}_j \tag{13}$$

where $\bar{C}_j$ is the average per-type capability. For $\rho < 1$, $C_{\text{eff}}$ is superlinear in $J$: each new specialization type adds more to the aggregate than the previous one (in terms of the diversity premium), even if the new type starts with the same per-type capability.

**Proposition 4** (Saturation Escape via Variety)**.** *Even with training saturation $h > 0$, the growth rate of $C_{eff}$ can remain positive if $J(t)$ is growing. Specifically, for constant per-type capability $C_j^*$ at the saturation ceiling:*

$$\frac{dC_{eff}}{dt} = \frac{1-\rho}{\rho} \cdot \frac{C_{eff}}{J} \cdot \dot{J} \cdot \left( \frac{(C_{J+1}^*)^{\rho}}{J^{-1} \sum_j (C_j^*)^{\rho}} \right) \tag{14}$$

*which is positive whenever a new type with $C_{J+1}^* > 0$ is created. The growth rate from variety expansion does not depend on $h$.*

*Proof.* Differentiating $C_{\text{eff}} = (\sum_j C_j^\rho)^{1/\rho}$ with respect to $J$, treating $J$ as continuous and using the chain rule:

$$\frac{\partial C_{\text{eff}}}{\partial J} = \frac{1}{\rho}\left(\sum_j C_j^\rho\right)^{1/\rho-1} \cdot (C_{J+1})^\rho \tag{15}$$

where $C_{J+1}$ is the capability of the newly created type. Since the saturation ceiling $C_j^* \leq 1/(h\delta)$ applies per type but not to the aggregate, variety expansion circumvents saturation. The aggregate grows as $J^{(1-\rho)/\rho}$ even when each individual $C_j$ is bounded. □

## 5. Layer 3: Diversity as Collapse Protection

The mesh's self-referential learning—agents training on data generated by other agents—risks model collapse. This section proves that the mesh's CES heterogeneity maintains training signal quality, connecting the production-theoretic parameter $\rho$ to an information-theoretic robustness condition.

### 5.1 The Model Collapse Framework

Following Shumailov et al. (2024), consider a generative model $Q_t$ trained on a mixture of authentic data from the true distribution $P$ (fraction $\alpha$) and synthetic data from the model's own previous generation $Q_{t-1}$ (fraction $1-\alpha$). The KL divergence from the true distribution evolves as:

$$\text{KL}(Q_{t+1}\|P) \geq \text{KL}(Q_t\|P) \quad \text{when } \alpha < \alpha_{\text{crit}} \tag{16}$$

The critical threshold $\alpha_{\text{crit}}$ depends on the model class. Below $\alpha_{\text{crit}}$, each generation amplifies the deviation from the true distribution: the model's outputs become progressively less diverse, tails of the distribution are truncated, and the model collapses toward a degenerate distribution. This is "model collapse"—a progressive degradation of capability through self-referential training.

For a single model training on its own outputs ($\alpha = 0$), collapse is inevitable. The outputs lack diversity because they are generated by a single distribution $Q_t$, which amplifies its own modes and suppresses its tails with each generation.

### 5.2 Effective Data Diversity in the Mesh

The mesh is not a single model. It is a collection of heterogeneous specialists whose outputs are drawn from different distributions. Agent $i$, specialized in task type $j^*(i)$, generates outputs from distribution $Q_i$ that reflects its specific training history, capability profile, and

specialization. From the perspective of agent $k \neq i$, training data from agent $i$ is not "self-generated"—it comes from a different distribution.

**Definition 5** (Effective External Data Fraction)**.** *For agent $i$ in a mesh with $J$ specialization types and CES parameter $\rho$, the effective external data fraction is:*

$$\alpha_{eff}(\rho, J) = \alpha_{ext} + (1 - \alpha_{ext}) \cdot D(\rho, J) \tag{17}$$

*where $\alpha_{ext}$ is the fraction of training data from sources outside the mesh (exogenous authentic data) and $D(\rho, J)$ is the diversity correction measuring the informational diversity of mesh-internal training data.*

The diversity correction $D(\rho, J)$ captures how different the mesh's internal training sources are from each agent's own outputs. Two agents with identical specializations ($Q_i = Q_k$) provide no diversity—training on each other's outputs is equivalent to self-training. Two agents with completely different specializations provide maximum diversity—their outputs are drawn from distributions as different as external data.

### 5.3 The Diversity Correction

Measure the pairwise distributional diversity between agents $i$ and $k$ by the Jensen-Shannon divergence $\text{JSD}(Q_i \| Q_k)$. For agents with CES-structured specializations, the expected pairwise divergence depends on $\rho$ and $J$.

**Lemma 2** (Specialization Implies Distributional Diversity)**.** *Let agents be fully specialized: agent $i$ produces outputs entirely from its specialization type $j^*(i)$. If the $J$ specialization types are distributed uniformly across the mesh, the expected diversity correction is:*

$$D(\rho, J) = \frac{J-1}{J} \cdot \left(1 - \rho^{1/(1-\rho)}\right) \tag{18}$$

*For $\rho < 1$ (complementary task types): $D > 0$, and $D$ is increasing in $J$ and decreasing in $\rho$. As $\rho \to 0$ (perfect complements): $D \to (J-1)/J$. As $\rho \to 1$ (perfect substitutes): $D \to 0$.*

*Proof.* When agents are fully specialized, agent $i$'s output distribution $Q_i$ has support concentrated on task type $j^*(i)$. For agent $k$ with $j^*(k) \neq j^*(i)$, the supports are disjoint, giving maximal Jensen-Shannon divergence. The fraction of other agents with different specializations is $(J-1)/J$ under uniform distribution. The term $(1-\rho^{1/(1-\rho)})$ quantifies how the CES substitution parameter translates production complementarity into distributional diversity: lower $\rho$ implies more differentiated outputs, hence greater informational diversity. At $\rho = 0$,

task types are perfectly complementary and outputs are maximally diverse; at $\rho = 1$, task types are perfect substitutes and outputs are identical, providing no diversity benefit. □

### 5.4 The Dual Role of $\rho$

**Theorem 1** (CES Heterogeneity as Collapse Protection). *Let $\alpha_{ext} \geq 0$ be the exogenous external data fraction and let $\rho < 1$ be the CES substitution parameter. The mesh avoids model collapse ($\alpha_{eff} > \alpha_{crit}$) whenever:*

$$\alpha_{ext} + (1 - \alpha_{ext}) \cdot \frac{J-1}{J} \cdot \left(1 - \rho^{1/(1-\rho)}\right) > \alpha_{crit} \tag{19}$$

*This condition can be satisfied even when $\alpha_{ext} < \alpha_{crit}$—that is, even when the mesh's external data supply is below the collapse threshold for any individual model—provided $J$ is sufficiently large and $\rho$ is sufficiently small.*

*Proof.* Substituting equation (**??**) into equation (**??**) gives condition (**??**) directly. The condition $\alpha_{\text{eff}} > \alpha_{\text{crit}}$ holds when:

$$D(\rho, J) > \frac{\alpha_{\text{crit}} - \alpha_{\text{ext}}}{1 - \alpha_{\text{ext}}} \tag{20}$$

The right side is positive (and the condition is binding) only when $\alpha_{\text{ext}} < \alpha_{\text{crit}}$—precisely when external data alone is insufficient to prevent collapse. The left side $D(\rho, J)$ increases as $\rho$ decreases and $J$ increases. For any $\alpha_{\text{crit}} \in (0, 1)$ and $\alpha_{\text{ext}} \in [0, \alpha_{\text{crit}})$, there exist $(\rho, J)$ pairs satisfying the condition.

Specifically, the minimum $J$ required for collapse avoidance is:

$$J_{\min}(\rho, \alpha_{\text{ext}}) = \left\lceil \frac{1}{1 - \frac{\alpha_{\text{crit}} - \alpha_{\text{ext}}}{(1-\alpha_{\text{ext}})(1-\rho^{1/(1-\rho)})}} \right\rceil \tag{21}$$

This is finite for $\rho < 1$ and decreasing in $\rho$ (more complementary specializations require fewer types to achieve collapse protection). □

**Remark 3** (One Parameter, Two Functions). *The CES parameter $\rho$ now does double duty. In the mesh paper, $\rho < 1$ generates the* diversity premium*: heterogeneous specialists collectively outperform a homogeneous centralized provider because the CES aggregate rewards breadth (Lemma 1 of Smirl 2026). In this paper, $\rho < 1$ generates* collapse protection*: heterogeneous specialists generate informationally diverse training data that prevents model collapse even when external data is scarce. The same force—agent heterogeneity, measured by $\rho$—drives both capability aggregation and informational robustness. This dual role is not a coincidence;*

*it reflects the structural identity between production complementarity and distributional diversity in the CES framework.*

## 6. The Central Theorem: Growth Regimes

We now unify the three layers into a complete characterization of the mesh's growth dynamics.

**Theorem 2** (Growth Regime Classification)**.** *Consider the mesh growth system defined by equations* (??), (??), (??), (??), *and* (??), *with the autocatalytic core existing for* $N > N_{auto}$ *(Proposition* ??*). The long-run behavior of* $C_{eff}(t)$ *falls into one of three regimes:*

***Regime (a): Convergence to a ceiling.*** *If* $\varphi_{eff} < 1$ *(equivalently,* $\beta_{\text{auto}} < (1-\varphi_0)/\varphi_0$*), training saturation* $h > 0$*, and variety is bounded (*$J \leq J_{\max} < \infty$*), then:*

$$C_{eff}(t) \to C_{max} \equiv J_{\max}^{(1-\rho)/\rho} \cdot \frac{1}{h\delta} \quad \text{as } t \to \infty \tag{22}$$

*The convergence rate is governed by* $1 - \varphi_{eff}$*. The ceiling* $C_{max}$ *is increasing in* $J_{\max}$ *(more specialization types raise the ceiling), decreasing in* $h$ *(faster saturation lowers it), and decreasing in* $\delta$ *(faster obsolescence lowers it). In this regime, the mesh's growth rate decelerates to zero, and the long-run growth rate of* $C_{eff}$ *equals the growth rate of the exogenous food set (frontier model releases).* This is the Baumol bottleneck.

***Regime (b): Balanced exponential growth.*** *If* $\varphi_{eff} = 1$ *(equivalently,* $\beta_{\text{auto}} = (1 - \varphi_0)/\varphi_0$*) and* $J(t)$ *grows endogenously at rate* $g_J > 0$*, then:*

$$C_{eff}(t) \sim C_{eff}(0) \cdot \exp\left[\left(\delta_g f^{\lambda} J_0^{\gamma_J} e^{\gamma_J g_J t} - \delta\right) t\right] \tag{23}$$

*The growth rate accelerates slowly as variety expands, but the dominant term is exponential. The Baumol bottleneck binds only if the exogenous input* $Z$ *grows slower than* $C$*, in which case the non-automated fraction* $(1 - \beta_{\text{auto}})$ *of the training process becomes the constraint and* $C_{eff}$ *growth decelerates to the exogenous rate. Even in the exponential regime, the long-run growth rate is bounded by* $g_Z/(1 - \beta_{\text{auto}})$ *where* $g_Z$ *is the growth rate of frontier model capability.*

***Regime (c): Finite-time singularity.*** *If* $\varphi_{eff} > 1$ *(equivalently,* $\beta_{\text{auto}} > (1 - \varphi_0)/\varphi_0$*),* $h = 0$ *(no training saturation), and* $\alpha_{eff} > \alpha_{crit}$ *(model collapse avoided), then:*

$$C_{eff}(t) = \left(C_0^{1-\Phi} - (1 - \Phi) \cdot \delta_g f^{\lambda} J^{\gamma_J} \cdot t\right)^{1/(1-\Phi)} \tag{24}$$

*where* $\Phi = \lambda + \varphi_{eff} - 1 > 1$. *This diverges at finite time:*

$$T_s = \frac{C_0^{1-\Phi}}{(\Phi - 1) \cdot \delta_g f^\lambda J^{\gamma_J}} \tag{25}$$

*The singularity requires three conditions simultaneously: (i)* $\varphi_{eff} > 1$*, which requires the mesh to automate more than* $(1 - \varphi_0)/\varphi_0$ *of its own training process; (ii)* $h = 0$*, no diminishing returns to training individual specialists; (iii)* $\alpha_{eff} > \alpha_{crit}$*, sufficient data diversity to avoid model collapse. The conjunction of these conditions is restrictive.*

*Proof. Regime (a):* With $\Phi < 1$, the growth rate $g_C = \delta_g f^\lambda C^{\Phi-1} J^{\gamma_J} - \delta$ is decreasing in $C$. The unique steady state $C^*$ satisfies $\delta_g f^\lambda (C^*)^{\Phi-1} J^{\gamma_J} = \delta$. With training saturation $h > 0$, each $C_j \leq 1/(h\delta)$ (Lemma **??**), and with bounded $J \leq J_{\max}$, the ceiling $C_{\max}$ is finite. Global stability follows from the monotone dynamical systems theorem (Hirsch 1985): the system is cooperative and bounded, hence converges to the unique equilibrium.

*Regime (b):* With $\Phi = 1$, the growth equation becomes $\dot{C} = \delta_g f^\lambda J(t)^{\gamma_J} - \delta C$, which is a linear ODE with time-varying coefficients. The solution is exponential with growth rate modulated by $J(t)$. If $J$ grows at rate $g_J$, then $C_{\text{eff}}$ grows at an accelerating exponential rate, bounded eventually by the Baumol constraint (Section 7): the non-automated fraction of training requires exogenous input $Z$ growing at rate $g_Z$, which bounds the long-run growth of $C$.

*Regime (c):* With $\Phi > 1$ and $h = 0$, the ODE $\dot{C} = \delta_g f^\lambda C^\Phi J^{\gamma_J}$ (ignoring depreciation, which is dominated) is a Bernoulli equation. The substitution $v = C^{1-\Phi}$ yields $\dot{v} = (1 - \Phi)\delta_g f^\lambda J^{\gamma_J}$, which integrates to $v(t) = v(0) + (1 - \Phi)\delta_g f^\lambda J^{\gamma_J} t$. Since $1 - \Phi < 0$, $v$ decreases linearly, reaching zero at $T_s$. Back-substituting gives $C(t) = v(t)^{1/(1-\Phi)} \to \infty$ as $t \to T_s$. □

Table 1: Growth regime classification.

| Regime | $\varphi_{\text{eff}}$ | $h$ | $J$ | Long-run $C_{\text{eff}}(t)$ |
|---|---|---|---|---|
| (a) Convergence | $< 1$ | $> 0$ | bounded | $\to C_{\max}$ (ceiling) |
| (b) Exponential | $= 1$ | $\geq 0$ | growing | $\sim e^{rt}$ |
| (c) Singularity | $> 1$ | $= 0$ | any | $\to \infty$ at $T_s < \infty$ |
| *Additional condition for all regimes:* $\alpha_{\text{eff}} > \alpha_{\text{crit}}$ (collapse avoided) | | | | |

**Remark 4** (Which Regime Is Likely?)**.** *The empirical evidence strongly favors regime (a) in the near term. Bloom et al. (2020) estimate* $\varphi \approx 0.5$–$0.7$ *for research productivity across multiple domains, implying* $\varphi_0 < 1$ *by a substantial margin. For* $\varphi_{eff}$ *to reach unity with* $\varphi_0 = 0.6$*, the autocatalytic fraction must reach* $\beta_{auto} = 0.67$*—the mesh must automate two-thirds of its own training improvement process. While not impossible, this requires training*

*agent capabilities substantially beyond current levels. Individual training interactions exhibit clear saturation ($h > 0$): repeated fine-tuning of the same model yields diminishing returns. And while variety expansion can escape per-type saturation, the total task space $J_{\max}$ is finite.*

*Regime (c) requires the conjunction of three conditions, each individually unlikely: $\varphi_{eff} > 1$ (near-complete training automation), $h = 0$ (no saturation), and $\alpha_{eff} > \alpha_{crit}$ (collapse avoided). The probability that all three hold simultaneously is small. The paper does not predict the singularity. It characterizes the conditions under which it would occur, and notes that these conditions are restrictive.*

*The most probable trajectory is: regime (a) with an increasing ceiling. As the mesh matures, $\beta_{auto}$ rises (pushing $\varphi_{eff}$ toward 1), $J$ expands (raising $C_{max}$), and the ceiling lifts—but never vanishes entirely, because the Baumol bottleneck (Section 7) anchors the long-run growth rate to exogenous frontier model improvement.*

## 7. The Baumol Bottleneck: Training Persistence as Rate Limiter

The mesh paper assumes training persistence: frontier model training remains centralized. This section derives training persistence as a *consequence* of the growth dynamics, not an assumption. The mechanism is Baumol's (1967) cost disease, applied to the mesh's internal production structure.

### 7.1 The Two-Sector Structure

Decompose the AI production process into two sectors:

**Sector 1: Inference and fine-tuning** (progressively automated). The mesh automates an increasing fraction $\beta(t)$ of inference and fine-tuning tasks. These are the tasks the mesh paper models: routing queries to specialists, generating responses, adapting models to new domains through fine-tuning. The mesh's productivity in this sector grows at rate $g_C$—the endogenous growth rate from the central model.

**Sector 2: Frontier model training** (non-automatable). Training frontier models from scratch requires tightly synchronized GPU clusters at scales of $10^{25}$+ FLOPs per run, with synchronization bandwidth as the binding constraint (Smirl 2026a, Section 4). This synchronization requirement is *topological*, not cost-based: it requires all-to-all communication within the training cluster, which distributed mesh architecture cannot provide. Frontier training productivity grows at exogenous rate $g_Z$, determined by centralized infrastructure investment.

### 7.2 Derivation of the Baumol Bottleneck

Following Aghion, Jones and Jones (2018), model the aggregate AI capability as a Cobb-Douglas composite of the two sectors:

$$C_{\text{eff}} = C_1^{\beta(t)} \cdot C_2^{1-\beta(t)} \tag{26}$$

where $C_1$ is the mesh's inference/fine-tuning capability (growing at $g_C$) and $C_2$ is frontier training capability (growing at exogenous rate $g_Z$). The automation fraction $\beta(t) \to 1$ as the mesh progressively automates more inference tasks.

The growth rate of the aggregate is:

$$g_{C_{\text{eff}}} = \beta(t) \cdot g_C + (1 - \beta(t)) \cdot g_Z + \dot{\beta}(t) \cdot \ln\left(\frac{C_1}{C_2}\right) \tag{27}$$

**Proposition 5** (Endogenous Baumol Bottleneck)**.** *As $\beta(t) \to 1$, the growth rate $g_{C_{eff}} \to g_Z$ regardless of $g_C$, provided $g_C > g_Z$ (the mesh's automated sector grows faster than the non-automated sector). The non-automatable sector becomes the binding constraint even as its share of total activity shrinks.*

*Proof.* Rewrite the growth rate using the cost share. In balanced growth, the cost share of sector 2 is $s_2 = (1-\beta) \cdot w_2 / [(1-\beta)w_2 + \beta w_1]$, where $w_i$ are the factor prices. When $g_C > g_Z$, the relative price of the non-automated sector rises: $w_2/w_1$ increases over time. The cost share $s_2$ rises even as the volume share $(1-\beta)$ falls—this is Baumol's cost disease.

In the limit $\beta \to 1$, the mesh has automated all automatable tasks. The only remaining input is frontier model training (sector 2). The aggregate growth rate satisfies:

$$\lim_{\beta \to 1} g_{C_{\text{eff}}} = g_Z + \lim_{\beta \to 1} \dot{\beta} \cdot \ln(C_1/C_2) \tag{28}$$

If $\dot{\beta} \to 0$ as $\beta \to 1$ (the last few tasks are hardest to automate—the topological constraint on training), then $g_{C_{\text{eff}}} \to g_Z$. The mesh's growth rate converges to the exogenous rate of frontier model improvement, regardless of how fast the mesh's internal productivity grows. □

### 7.3 Closing the Circle

The Baumol bottleneck connects Paper 4 back to Smirl (2026a). The chain of determination is:

(i) *Concentrated investment* (Smirl 2026a): Datacenter capital investment, driven by the

power-law scaling relationship between model size and performance, finances the GPU clusters that train frontier models. The rate of frontier model improvement $g_Z$ is determined by the rate of investment.

(ii) *Learning curves* (Smirl 2026a): The same concentrated investment finances the 3D memory stacking and advanced packaging learning curves ($\alpha = 0.23$) that enable distributed inference. The crossing point $x(t) = 0$ is when distributed architecture becomes cost-competitive.

(iii) *Mesh formation* (Smirl 2026): After crossing, the mesh self-organizes into a heterogeneous specialized network whose collective capability exceeds centralized provision at $N > N^*$.

(iv) *Endogenous growth* (this paper): The mesh improves itself through autocatalytic training, self-referential learning, and variety expansion. The growth regime depends on $\varphi_{\text{eff}}$, $h$, and $\alpha$.

(v) *Baumol ceiling* (this paper): The mesh's growth rate converges to $g_Z$—the rate of frontier model improvement—which is determined by the concentrated investment of step (i).

The circle closes. The concentrated capital that creates the crossing also determines the ceiling. The mesh amplifies frontier model improvement but cannot exceed it indefinitely. This is a falsifiable prediction: the mesh's capability growth rate should correlate with, and be bounded by, the rate of frontier model releases from centralized providers (Prediction 4, Section 10).

The five-step chain above instantiates the general $N_{\text{eff}}$-level hierarchy formalized in Smirl (2026d, Theorem 4.3). The current-era hierarchy depth $N_{\text{eff}} = 4$ reflects four distinct timescale gaps (hardware learning $\gg$ mesh adoption $\gg$ capability growth $\gg$ settlement dynamics); the companion paper proves that this depth is endogenously determined by the technology's spectral gap structure and will itself evolve as the AI transition matures. The Baumol bottleneck—this paper's central result—is the mathematical expression of the hierarchical ceiling: level $n$'s growth rate is bounded by level $n-1$'s rate for any $N_{\text{eff}} \geq 2$.

## 8. Connection to Settlement Infrastructure

The mesh paper (Smirl 2026, Section 8) derives that the mesh's routing and compensation requirements endogenously generate the need for a programmable settlement layer. The autocatalytic growth dynamics of this paper intensify that requirement.

In the mesh paper's static analysis, the settlement layer processes $O(N\langle k\rangle)$ inference transactions per second between end users and specialists. With endogenous capability growth, the settlement layer must additionally process the micro-transactions of the autocatalytic loop:

(i) *Training agent compensation:* Agents that improve other agents must be compensated. Each training operation in the RAF set (Section 3) involves a training agent investing compute to improve a target agent. Without compensation, the training agent has no incentive to allocate resources to improvement rather than inference.

(ii) *Data marketplace transactions:* Self-referential learning (Mechanism 2) requires agents to share operational data—queries, responses, evaluations—as training inputs. This data has value, and efficient allocation requires a price mechanism.

(iii) *Variety expansion incentives:* Creating new specialization types (Mechanism 3) is a costly innovation requiring dedicated training compute. The innovating agent must capture returns sufficient to cover the investment, which requires a mechanism for pricing access to the new specialization.

The transaction volume from autocatalytic operations scales as $O(|\mathcal{R}| \cdot f \cdot N)$, where $|\mathcal{R}|$ is the size of the RAF set and $f$ is the training allocation fraction. As the autocatalytic core expands (via the Jain-Krishna process of Section 3.5), $|\mathcal{R}|$ grows, accelerating the settlement layer requirement.

This connects to Smirl (2026b, "The Monetary Productivity Gap"): the settlement layer is not merely needed for routing compensation—it is needed for the micro-transactions that the autocatalytic loop requires. The mesh's capability growth rate may be constrained by settlement infrastructure before it is constrained by any of the three parameters ($\varphi_{\text{eff}}$, $h$, $\alpha$) analyzed in the central model. In this case, the binding constraint on mesh improvement is *monetary*, not technological—a conclusion that reinforces the monetary productivity gap analysis.

## 9. Frameworks Considered and Rejected

Several candidate frameworks were evaluated for the formal model and rejected for specific technical reasons. This section documents the evaluation to clarify modeling choices.

**Eigen's Hypercycle (Eigen & Schuster 1977).** The hypercycle model describes a cyclic network of autocatalytic replicators where each species catalyzes the replication of the next. The autocatalytic structure is conceptually correct for the mesh—agents catalyze each

other's improvement. However, the hypercycle imposes a conservation law: $\sum_i x_i = \text{const}$ (total concentration is fixed). This forces zero-sum dynamics among capability types. The mesh has no such conservation law—it is an open system where total capability can grow. The RAF framework (Hordijk & Steel 2004) provides the autocatalytic structure without the conservation constraint.

**Chemical Reaction Network Theory (Feinberg 2019).** CRNT provides powerful results on the existence and uniqueness of equilibria in reaction networks, particularly the deficiency zero theorem: networks with deficiency zero have a unique positive equilibrium that is locally asymptotically stable within each stoichiometric compatibility class. However, CRNT assumes closed systems with stoichiometric conservation laws. The mesh is open (it receives exogenous base models and produces capability improvements that are not conserved in a stoichiometric sense). Moreover, CRNT characterizes equilibrium existence, not growth trajectories—and the central question of this paper is the growth trajectory.

**NK Fitness Landscapes (Kauffman 1993).** The NK model of rugged fitness landscapes with epistatic interactions is conceptually appropriate for co-evolutionary dynamics among mesh agents. However, the NK framework lacks analytical results on convergence or divergence of the co-evolutionary process. The co-evolutionary extension (NKC model with multiple interacting landscapes) is an open problem in complexity science: it is not known whether the dynamics converge, cycle, or exhibit chaotic behavior. Useful as motivation for the variety expansion mechanism, but not as formal machinery for growth regime characterization.

**Spin Glasses (Edwards-Anderson 1975; Sherrington-Kirkpatrick 1975).** Rejected for the same reason as in the mesh paper (Smirl 2026, Section 9): spin glass models require frustrated interactions (a mix of positive and negative couplings). In the mesh, all training interactions are positive—being improved by another agent is always (weakly) beneficial. There is no frustration. The Lotka-Volterra mutualistic framework (Bastolla et al. 2009) correctly captures the all-positive interaction structure with saturation.

## 10. Falsifiable Predictions

The model generates six predictions extending the mesh paper's prediction set. Each has timing, observable consequences, and failure conditions.

**Prediction 1: Autocatalytic Threshold Timing.** The mesh achieves self-sustaining capability improvement (a RAF set forms within the mesh's active agents) within 3 years of crystallization (Prediction 1 of Smirl 2026). Observable as: mesh capability improving on standard benchmarks without new base model releases from centralized providers. Specifi-

cally, the mesh's average score on established benchmarks increases by $\geq 5\%$ during a period of $\geq 6$ months with no major frontier model release. *Timing:* 2033–2036. *Evidence against:* mesh capability plateauing or declining during any 12-month period without new base model releases through 2038.

**Prediction 2: Training Agent Emergence.** Specialized training agents—agents whose primary function is improving other agents rather than serving end-user queries—emerge within the mesh and capture $> 10\%$ of internal mesh transactions within 5 years of crystallization. Observable in: the composition of mesh transaction types shifting from pure inference toward training, evaluation, and fine-tuning operations. *Timing:* 2035–2038. *Evidence against:* mesh transactions remaining $> 95\%$ pure inference through 2040, with no significant training agent specialization.

**Prediction 3: Diversity-Collapse Protection.** Heterogeneous meshes (high $J$, low $\rho$) maintain or improve capability when training on internally generated data, while homogeneous networks (low $J$, high $\rho$) exhibit model collapse. Observable in: benchmark performance over time for mesh configurations with varying diversity levels. Specifically, meshes with $J \geq 10$ and estimated $\rho \leq 0.5$ maintain stable or improving benchmark scores when trained on $> 50\%$ synthetic internal data, while homogeneous networks ($J \leq 3$) show degradation under the same conditions. *Evidence against:* homogeneous networks showing no degradation from synthetic data training, or heterogeneous meshes degrading despite high diversity.

**Prediction 4: Baumol Bottleneck Binding.** The mesh's capability growth rate correlates with, and is bounded by, the rate of frontier model releases from centralized providers. Observable as: (i) mesh capability plateauing between major model releases and jumping after them; (ii) the ratio of mesh capability growth to frontier model improvement converging to a constant multiplier. Quantitatively, the mesh's annualized capability growth rate (measured by benchmark improvement) should track within $1.5\times$ of the annualized frontier model improvement rate. *Timing:* 2034–2040. *Evidence against:* mesh capability growth rate exceeding $3\times$ the frontier model release rate sustained over $> 2$ years, indicating the Baumol bottleneck is not binding.

**Prediction 5: $\varphi_{\text{eff}}$ Estimate.** The mesh's effective training productivity elasticity is measurable from capability benchmarks and training compute data. Based on the empirical evidence for $\varphi_0$ (Bloom et al. 2020) and expected $\beta_{\text{auto}}$ trajectories, the initial $\varphi_{\text{eff}}$ should be 0.6–0.8 (regime (a), converging), potentially rising toward 0.9–1.0 as the autocatalytic core matures. Observable as: the elasticity of benchmark improvement with respect to training compute invested, measured across mesh capability generations. *Evidence against:* $\varphi_{\text{eff}} > 1.0$ sustained over $> 1$ year (indicating regime (c), which the model identifies as unlikely).

**Prediction 6: Variety Expansion Rate.** The number of effective specialization types $J$ grows at a rate determined by unmet demand signals and the RAF existence threshold (equation **??**). Observable in: the diversity of fine-tuned models available on the mesh, measured by the effective number of distinct capability clusters. Quantitatively, $J$ should grow at 15–30% annually during the rapid growth phase, decelerating as $J$ approaches $J_{\max}$. *Evidence against:* $J$ remaining constant or declining after crystallization, indicating no endogenous variety expansion.

## 11. Conclusion

This paper has characterized what happens when the mesh can improve itself. The mesh paper's fixed-capability assumption—each agent's capability redistributes but does not grow—is replaced with endogenous capability dynamics driven by three mechanisms: autocatalytic training, self-referential learning, and variety expansion.

The answer to the ceiling question is regime-dependent. Three parameters govern the growth dynamics. The training productivity elasticity $\varphi$ determines whether capability growth decelerates, is constant, or accelerates. Training saturation $h$ determines whether individual improvement interactions have diminishing returns. The external data fraction $\alpha$, maintained above the model collapse threshold by the mesh's CES heterogeneity, determines whether self-referential learning is productive.

Three results distinguish this analysis from the existing literature on AI and economic growth. First, the autocatalytic existence threshold $N_{\text{auto}}$ establishes that self-sustaining capability improvement emerges at a mesh size scaling logarithmically with system complexity—the same combinatorial logic that makes the mesh paper's critical mass $N^*$ decrease with diversity. The RAF framework from origin-of-life theory provides the formal structure; the Jain-Krishna adaptive network dynamics predict the temporal pattern of reorganization cascades within the autocatalytic core.

Second, the effective training productivity elasticity $\varphi_{\text{eff}}$ is derived from the mesh's microstructure. Individual training interactions have $\varphi_0 < 1$ (Bloom et al. 2020). But autocatalytic coupling—training agents that improve other training agents—amplifies the effective elasticity: $\varphi_{\text{eff}} = \varphi_0/(1-\beta_{\text{auto}}\varphi_0)$, which exceeds $\varphi_0$ for any positive automation fraction $\beta_{\text{auto}}$. Whether $\varphi_{\text{eff}}$ reaches unity—the knife-edge between convergence and exponential growth—depends on whether the mesh can automate a sufficient fraction of its own improvement process before the Baumol bottleneck binds.

Third, the CES parameter $\rho$ does double duty: it governs both the diversity premium for capability aggregation (the mesh paper's result) and the diversity protection against model

collapse (this paper's result). The same heterogeneity that makes the mesh capable also makes it robust.

The Baumol bottleneck emerges from the growth dynamics rather than being imposed as an assumption. As the mesh automates progressively more tasks, the non-automatable sector—frontier model training—becomes the binding constraint. The mesh's growth rate converges to the exogenous rate of frontier model improvement. The concentrated infrastructure investment modeled in Smirl (2026a) determines both when the crossing occurs and the ceiling the mesh approaches. The circle closes.

The most likely near-term trajectory is regime (a): convergence to a ceiling that rises as the autocatalytic core matures, variety expands, and frontier models improve. The mesh is a multiplier, not a generator. Whether it transitions to regime (b)—sustained exponential growth—depends on the empirical trajectory of $\beta_{\text{auto}}(t)$ and $J(t)$, which are measurable quantities (Predictions 5 and 6). The singularity regime (c) requires conditions that are each individually demanding and collectively unlikely to hold simultaneously.

The organizational form that emerges is not merely a static division of labor, as the mesh paper characterizes, but a dynamical system capable of self-improvement. The mesh does not just distribute inference—it creates a substrate for the endogenous growth of intelligence. The rate of that growth is the empirical question this paper has formalized.

The companion paper (Smirl 2026d) generalizes the hierarchy: the four-level structure is the current-era application of an $N_{\text{eff}}$-level framework whose depth is endogenously determined. As the autocatalytic mesh matures, timescale convergence may reduce $N_{\text{eff}}$—for instance, if training and capability growth merge onto a single timescale—while novel processes may crystallize new levels. The Baumol bottleneck and the $R_0$ activation threshold are structural features of any $N_{\text{eff}} \geq 2$ hierarchy, not artifacts of the specific four-level decomposition.

# A. RAF Theory and Supporting Results

## A.1 Background on RAF Sets

The Reflexively Autocatalytic and Food-generated (RAF) framework was introduced by Hordijk and Steel (2004) to formalize Kauffman's (1971, 1993) theory of autocatalytic sets in the context of the origin of life. The original question was: in a random soup of molecular species, how large must the system be for a self-sustaining network of catalyzed reactions to emerge? The answer—that the threshold scales logarithmically with system complexity—is one of the foundational results in origin-of-life theory.

The formal setup: let $X$ be a set of molecule types, $\mathcal{R}$ a set of reactions (each with

reactants, products, and potential catalysts), and $F \subset X$ a food set of molecules available exogenously. A RAF set is a subset $\mathcal{R}' \subseteq \mathcal{R}$ that is simultaneously autocatalytic (every reaction is catalyzed by a product of the set or the food set) and food-generated (every reactant can be built from $F$ through reactions in $\mathcal{R}'$).

The key theorem (Hordijk & Steel 2004): in a random catalytic reaction system with $n$ molecule types and catalysis probability $p$ per type-reaction pair, a RAF set exists with high probability when $p > 1/n$. Since $p$ increases with the number of molecules in the system (each molecule provides an additional potential catalyst), and the threshold scales as $1/n$ which *decreases* with system size, the RAF existence threshold is remarkably low.

The translation to the mesh is direct. "Molecule types" are capability types ($J$). "Reactions" are training operations. "Catalysts" are training agents with the requisite capability to direct a training operation. "Food set" is base model capabilities from centralized training. The Hordijk-Steel result translates to Proposition **??**: the autocatalytic threshold $N_{\text{auto}}$ scales logarithmically with system complexity.

### A.2 Proof Details for Proposition ??

The derivation of $\varphi_{\text{eff}}$ follows the Aghion-Jones-Jones (2018) framework applied to the mesh's internal structure. Consider the training improvement process as consisting of a unit continuum of subtasks $s \in [0, 1]$. Each subtask $s$ requires a quality-adjusted input $x(s)$ to produce its contribution to capability improvement.

For automated subtasks ($s \in [0, \beta_{\text{auto}}]$), the input is supplied by mesh agents with capability $C$:

$$x(s) = A_s \cdot C^{\varphi_0}, \quad s \in [0, \beta_{\text{auto}}] \tag{29}$$

For non-automated subtasks ($s \in (\beta_{\text{auto}}, 1]$), the input is exogenous:

$$x(s) = Z, \quad s \in (\beta_{\text{auto}}, 1] \tag{30}$$

The aggregate improvement is Cobb-Douglas across subtasks:

$$\dot{C} \propto \exp\left(\int_0^1 \ln x(s)\, ds\right) = \left(\prod_{s \leq \beta_{\text{auto}}} A_s\right) \cdot C^{\beta_{\text{auto}}\varphi_0} \cdot Z^{1-\beta_{\text{auto}}} \tag{31}$$

The growth rate of $C$ is:

$$g_C = \beta_{\text{auto}}\varphi_0 \cdot g_C + (1 - \beta_{\text{auto}}) g_Z + \text{const} \tag{32}$$

Solving: $g_C(1 - \beta_{\text{auto}}\varphi_0) = (1 - \beta_{\text{auto}})g_Z + \text{const}$, so:

$$g_C = \frac{(1 - \beta_{\text{auto}})g_Z}{1 - \beta_{\text{auto}}\varphi_0} + \text{const} \tag{33}$$

The effective elasticity of $C$ with respect to itself in the production of $\dot{C}$ is $\varphi_{\text{eff}} = \beta_{\text{auto}}\varphi_0/(1 - \beta_{\text{auto}}\varphi_0 + \beta_{\text{auto}}\varphi_0) = \varphi_0/(1 - \beta_{\text{auto}}\varphi_0)$, confirming equation (**??**).

### A.3 Weitzman Recombinant Growth Connection

Weitzman (1998) models the growth of ideas as a combinatorial process: new ideas are produced by recombining existing ideas, and the number of potential recombinations grows faster than the number of existing ideas. This produces growth rates that accelerate over time.

The mesh's variety expansion mechanism (Section 4.6) has a Weitzman interpretation. New specialization types are produced by combining existing specializations: a medical-legal specialist combines medical reasoning and legal analysis capabilities. The number of potential combinations grows as $\binom{J}{2} \sim J^2/2$, so the potential for variety expansion accelerates with existing variety. Equation (**??**) is a simplified reduced form; the Weitzman recombinant structure provides a microfoundation for why $\dot{J}$ can be superlinear in $J$ over portions of the trajectory, further supporting the saturation escape mechanism of Proposition **??**.

### A.4 Nordhaus Singularity Analysis

Nordhaus (2021) asks whether we are approaching an economic singularity—a regime in which economic growth becomes superexponential. His analysis identifies conditions under which standard growth models produce accelerating growth, and concludes that current empirical trends do not support the singularity hypothesis.

This paper's regime (c) is precisely Nordhaus's singularity condition applied to the mesh's internal dynamics. The contribution is identifying the three specific parameters ($\varphi_{\text{eff}}$, $h$, $\alpha$) whose conjunction determines whether the singularity obtains for the mesh. The paper's conclusion aligns with Nordhaus: the conditions are restrictive and unlikely to hold simultaneously. The mesh's most probable trajectory is regime (a)—convergence to a ceiling—with the Baumol bottleneck as the rate limiter.

## References


[1] Aghion, P., Jones, B. F., & Jones, C. I. (2018). Artificial intelligence and economic growth. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The Economics of Artificial Intelligence* (pp. 237–282). University of Chicago Press.

[2] Bastolla, U., Lässig, M., Manrubia, S. C., & Valleriani, A. (2009). Biodiversity in model ecosystems, I: Coexistence conditions for competing species. *Journal of Theoretical Biology*, 235(4), 521–530.

[3] Baumol, W. J. (1967). Macroeconomics of unbalanced growth: The anatomy of urban crisis. *American Economic Review*, 57(3), 415–426.

[4] Becker, G. S., & Murphy, K. M. (1992). The division of labor, coordination costs, and knowledge. *Quarterly Journal of Economics*, 107(4), 1137–1160.

[5] Bianconi, G., & Barabási, A.-L. (2001). Bose-Einstein condensation in complex networks. *Physical Review Letters*, 86(24), 5632–5635.

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

[7] Edwards, S. F., & Anderson, P. W. (1975). Theory of spin glasses. *Journal of Physics F: Metal Physics*, 5(5), 965–974.

[8] Eigen, M., & Schuster, P. (1977). The hypercycle: A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften*, 64(11), 541–565.

[9] Feinberg, M. (2019). *Foundations of Chemical Reaction Network Theory*. Springer.

[10] Hirsch, M. W. (1985). Systems of differential equations that are competitive or cooperative. II: Convergence almost everywhere. *SIAM Journal on Mathematical Analysis*, 16(3), 423–439.

[11] Hofbauer, J., & Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.

[12] Hordijk, W., & Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4), 451–461.

[13] Jain, S., & Krishna, S. (1998). Autocatalytic sets and the growth of complexity in an evolutionary model. *Physical Review Letters*, 81(25), 5684–5687.

[14] Jain, S., & Krishna, S. (2001). A model for the emergence of cooperation, interdependence, and structure in evolving networks. *Proceedings of the National Academy of Sciences*, 98(2), 543–547.

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

[16] Jones, C. I. (2005). Growth and ideas. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1B, pp. 1063–1111). Elsevier.

[17] Kauffman, S. A. (1971). Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems. *Journal of Cybernetics*, 1(1), 71–96.

[18] Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.

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

[20] Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5), S71–S102.

[21] Sherrington, D., & Kirkpatrick, S. (1975). Solvable model of a spin-glass. *Physical Review Letters*, 35(26), 1792–1796.

[22] Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, 631, 755–759.

[23] Smirl, C. (2026a). Endogenous decentralization: How concentrated capital investment finances the learning curves that enable distributed alternatives. Working paper, Tufts University.

[24] Smirl, C. (2026). The mesh equilibrium: How heterogeneous specialized agents self-organize to exceed centralized provision after the crossing point. Working paper, Tufts University.

[25] Smirl, C. (2026b). The monetary productivity gap. Working paper, Tufts University. Forthcoming.

[26] Smirl, C. (2026d). Complementary heterogeneity: How CES curvature unifies the four-level hierarchy. Working paper, Tufts University.

[27] Weitzman, M. L. (1998). Recombinant growth. *Quarterly Journal of Economics*, 113(2), 331–360.