Structural Transformation of Innovation (CLM2)

Detailed Derivations of the Model and Quantification

Martí Mestieri

Environment: Intermediate varieties, Bertrand, and markups

Each sector i has a unit continuum of varieties $v \in [0, 1]$. Output aggregator:

$$Y_i(t) = \exp\left(\int_0^1 \log X_{iv}(t)\,\mathrm{d}v\right).$$

Intermediate technology (frontier firm):

$$X_{iv}(t) = Q_{iv}(t)\,L_{iv}(t).$$

Competitive fringe knows a vintage with productivity Q_{iv}/χ. Bertrand $\Rightarrow$ frontier price:

$$P_{iv}(t) = \frac{\chi}{Q_{iv}(t)}.$$

Implication: baseline markup is degenerate at χ (no dispersion).
(You can discuss extensions with partial step sizes later.)

Markup distribution (baseline and a teaching extension)

Baseline in the paper: $\mu_{iv} = \chi$ for all i, v, t (degenerate distribution).
Reason: fringe always at Q/χ, and frontier always sets limit price χ/Q.

Extension for class discussion: If entrants sometimes improve by a factor $\Delta \equiv Q_{\text{new}}/Q_{\text{old}}$ and fringe is at Q_{old}, then limit pricing gives $\mu = \min\{\chi, \Delta\}$ and markup dispersion comes from the distribution of Δ.
Key object to compute in that extension:

$$\mathbb{P}(\mu = \chi) = \mathbb{P}(\Delta \geq \chi).$$

(Then use the innovation step-size distribution to get $\mathbb{P}(\Delta \geq \chi)$.)

Fréchet recap (the workhorse)

A Fréchet CDF with shape $\varsigma > 0$ and scale $K > 0$:

$$F(Q) = \exp\left(-KQ^{-\varsigma}\right), \qquad Q > 0.$$

PDF:

$$f(Q) = K\varsigma Q^{-(1+\varsigma)} \exp\left(-KQ^{-\varsigma}\right).$$

Tail:

$$\mathbb{P}(Q > x) = 1 - F(x) \sim Kx^{-\varsigma} \quad \text{for large } x.$$

First and second moments (levels)

For $m < \varsigma$,

$$\mathbb{E}[Q^m] = K^{-m/\varsigma}\,\Gamma\left(1 - \frac{m}{\varsigma}\right).$$

Hence (when they exist):

$$\mathbb{E}[Q] = K^{-1/\varsigma}\Gamma\left(1 - \frac{1}{\varsigma}\right), \qquad \mathbb{E}[Q^2] = K^{-2/\varsigma}\Gamma\left(1 - \frac{2}{\varsigma}\right).$$

Variance (requires $\varsigma > 2$):

$$\mathrm{Var}(Q) = K^{-2/\varsigma}\left[\Gamma\left(1 - \frac{2}{\varsigma}\right) - \Gamma\left(1 - \frac{1}{\varsigma}\right)^2\right].$$

Most applications take ς big enough so moments exist.

Log statistics (often more stable)

Let $Y = \log Q$. Then $Q = e^Y$ and

$$\mathbb{P}(Y \leq y) = \exp\left(-Ke^{-\varsigma y}\right).$$

This is a Gumbel-type form.
Useful identity: If $U \equiv KQ^{-\varsigma}$, then $U \sim \mathrm{Exp}(1)$.
So:

$$\log Q = \frac{1}{\varsigma}\left(\log K - \log U\right).$$

Hence:

$$\mathbb{E}[\log Q] = \frac{1}{\varsigma}(\log K + \gamma_E), \quad \mathrm{Var}(\log Q) = \frac{\pi^2}{6\varsigma^2},$$

where γ_E is Euler's constant.

Order statistics: max of Fréchets is Fréchet (max stability)

Let $Q_1, \ldots, Q_n$ i.i.d. with CDF $F(Q) = \exp(-KQ^{-\varsigma})$. Then the maximum $M = \max_k Q_k$ satisfies:

$$\mathbb{P}(M \leq Q) = \prod_{k=1}^{n} F(Q) = \exp\left(-(nK)Q^{-\varsigma}\right).$$

So $M \sim \text{Frećhet}(\text{scale} = nK, \text{shape} = \varsigma)$.
Key punchline: maxima update the *scale* parameter, keep the *shape*. This is why the model can track dynamics by tracking $K_i(t)$ only.

Spillovers as a max across source sectors

When innovating/adopting in destination sector i, a researcher draws one candidate prior idea from each source sector j:

$$\tilde{Q}_{oj} \sim \text{Fréchet}(K_j(t), \varsigma).$$

Applicability discount/cost $\phi_{ij} \in (0, 1]$ maps source quality to usable quality:

$$Q_{oj}^{(i)} = \frac{\tilde{Q}_{oj}}{\phi_{ij}}.$$

The best usable prior idea is the max:

$$Q_o = \max_j Q_{oj}^{(i)} = \max_j \left\{ \frac{\tilde{Q}_{oj}}{\phi_{ij}} \right\}.$$

Distribution of the best usable idea: compute the CDF

Compute $\mathbb{P}(Q_o \leq q)$:

$$\mathbb{P}\left(\max_j \frac{\tilde{Q}_{oj}}{\phi_{ij}} \leq q\right) = \prod_j \mathbb{P}(\tilde{Q}_{oj} \leq \phi_{ij} q).$$

Since $\tilde{Q}_{oj}$ is Fréchet:

$$\mathbb{P}(\tilde{Q}_{oj} \leq \phi_{ij} q) = \exp\left(-K_j(t)(\phi_{ij} q)^{-\varsigma}\right) = \exp\left(-K_j(t)\phi_{ij}^{-\varsigma} q^{-\varsigma}\right).$$

Multiply across j:

$$\mathbb{P}(Q_o \leq q) = \exp\left(-\left[\sum_j K_j(t)\phi_{ij}^{-\varsigma}\right] q^{-\varsigma}\right).$$

Define spillover stock:

$$S_i(t) \equiv \sum_j \phi_{ij}^{-\varsigma} K_j(t).$$

Then:

$$\tilde{F}_i(q,t) = \exp\left(-S_i(t)\, q^{-\varsigma}\right).$$

Innovation arrival and the forward equation (setup)

In each sector i, R&D labor is $Z_i(t)$. Innovations arrive as a Poisson process with intensity proportional to $Z_i(t)^{1-\phi}$.
Each innovation proposes a new frontier quality in a variety:

$$Q_{\text{new}} = Q_n \cdot Q_o^{\beta_i},$$

where:

- Q_n is a "new idea draw" (Pareto tail / Fréchet-compatible),
- Q_o is the best usable prior idea in sector i (distribution $\tilde{F}_i$ above),
- $\beta_i \in (0, 1)$ is the intertemporal spillover elasticity.

A variety frontier updates if $Q_{\text{new}} > Q$ (current frontier). We want the law of motion for $F_i(Q, t) \equiv \mathbb{P}(Q_{iv}(t) \leq Q)$.

Note on Poisson processes in economic growth (general)*

A Poisson process $\{N(t)\}_{t\geq 0}$ with (possibly time-varying) intensity $\lambda(t)$ satisfies:

- **Independent increments:**

$$N(t+s) - N(t) \ \perp \ N(t).$$

- **Small-time behavior:**

$$\Pr(N(t+\Delta t) - N(t) = 1) = \lambda(t)\Delta t + o(\Delta t),$$

$$\Pr(N(t+\Delta t) - N(t) \geq 2) = o(\Delta t).$$

- **Expected arrivals:**

$$\mathbb{E}[N(t)] = \int_0^t \lambda(s)\, ds.$$

- **Memorylessness:** conditional on no event yet, the remaining waiting time is exponential.

Why growth models use Poisson processes:

- smooth aggregation of discrete innovations
- tractable law of motion for distributions
- clear mapping between flow rates and hazards

Poisson processes in this model (innovation and replacement)

In this paper, Poisson processes govern **innovation attempts** and **frontier replacement**.

- Innovation attempts in sector i arrive with intensity:

$$\lambda_i(t) = \eta_i Z_i(t)^{1-\phi}.$$

- Each attempt draws a candidate productivity Q_{new}. A replacement occurs if:

$$Q_{\text{new}} > Q \quad \text{(current frontier)}.$$

- Replacement hazard for a variety with frontier Q:

$$h_i(Q,t) = \lambda_i(t)\,\Pr(Q_{\text{new}} > Q) \;\propto\; \lambda_i(t)\,Q^{-\varsigma}.$$

- This hazard representation implies the Kolmogorov forward equation:

$$\partial_t \log F_i(Q,t) = -h_i(Q,t),$$

  which preserves the Fréchet form of the distribution.

Economic intuition:

- Poisson arrivals smooth innovation over time
- High-productivity frontiers are harder to displace
- Aggregate growth emerges from many independent micro jumps

Kolmogorov forward equation: hazard representation

Fix a threshold $Q > 0$. $F_i(Q, t)$ increases when varieties with frontier above Q are "pushed down" below Q (doesn't happen), and decreases when varieties with frontier $\leq Q$ get upgraded above Q (does happen).
So the forward equation is:

$$\partial_t F_i(Q, t) = -\lambda_i(t)\, \mathbb{P}(\text{upgrade crosses } Q \mid Q_{iv}(t) \leq Q)\, F_i(Q, t),$$

where $\lambda_i(t) \propto Z_i(t)^{1-\phi}$ is the arrival rate of innovation attempts per variety.
Equivalently, in log form:

$$\partial_t \log F_i(Q, t) = -\lambda_i(t)\, \mathbb{P}(Q_{\text{new}} > Q \mid Q_{iv}(t) \leq Q).$$

This is the "hazard on the CDF" logic used to get equation (72) in the Appendix.

Compute the crossing probability: conditioning and integration

We need $\mathbb{P}(Q_{\text{new}} > Q)$:

$$\mathbb{P}(Q_n Q_o^{\beta_i} > Q) = \int_0^\infty \mathbb{P}\left(Q_n > \frac{Q}{q_o^{\beta_i}}\right) \tilde{f}_i(q_o, t)\, \mathrm{d}q_o.$$

Assume Pareto tail for Q_n consistent with Fréchet scaling:

$$\mathbb{P}(Q_n > x) = \tilde{\eta}_i\, x^{-\varsigma} \quad \text{(up to constants)}.$$

Then:

$$\mathbb{P}\left(Q_n > \frac{Q}{q_o^{\beta_i}}\right) = \tilde{\eta}_i \left(\frac{Q}{q_o^{\beta_i}}\right)^{-\varsigma} = \tilde{\eta}_i\, Q^{-\varsigma}\, q_o^{\beta_i \varsigma}.$$

Plug in:

$$\mathbb{P}(Q_{\text{new}} > Q) = \tilde{\eta}_i\, Q^{-\varsigma} \int_0^\infty q_o^{\beta_i \varsigma} \tilde{f}_i(q_o, t)\, \mathrm{d}q_o.$$

Solve the integral explicitly (the key step)

We have $\tilde{f}_i(\cdot, t)$ is Fréchet with scale $S_i(t)$ and shape ς:

$$\tilde{F}_i(q, t) = \exp(-S_i(t)q^{-\varsigma}).$$

So the moment formula gives, for $m = \beta_i\varsigma < \varsigma$:

$$\int_0^\infty q^{\beta_i\varsigma}\tilde{f}_i(q, t)\,\mathrm{d}q = \mathbb{E}[q^{\beta_i\varsigma}] = S_i(t)^{\beta_i}\,\Gamma(1 - \beta_i).$$

(Here we used $\mathbb{E}[q^m] = S^{-m/\varsigma}\Gamma(1 - m/\varsigma)$ and set $m = \beta_i\varsigma$.)
Therefore:

$$\mathbb{P}(Q_{\text{new}} > Q) = \tilde{\eta}_i\,\Gamma(1 - \beta_i)\,S_i(t)^{\beta_i}\,Q^{-\varsigma}.$$

Close the forward equation and recover Fréchet preservation

Put it back into:

$$\partial_t \log F_i(Q,t) = -\lambda_i(t)\mathbb{P}(Q_{\text{new}} > Q).$$

Using $\lambda_i(t) \propto Z_i(t)^{1-\phi}$:

$$\partial_t \log F_i(Q,t) = -\underbrace{[\tilde{\eta}_i \Gamma(1-\beta_i)]}_{\text{constant}} Z_i(t)^{1-\phi}\, S_i(t)^{\beta_i}\, Q^{-\varsigma}.$$

This has the form:

$$\partial_t \log F_i(Q,t) = -\dot{K}_i(t)\, Q^{-\varsigma}.$$

Integrate over time:

$$\log F_i(Q,t) = -K_i(t)Q^{-\varsigma}, \qquad F_i(Q,t) = \exp(-K_i(t)Q^{-\varsigma}).$$

Thus the distribution stays Fréchet, and all dynamics are in $K_i(t)$.

Law of motion for $K_i(t)$

From the coefficient mapping above:

$$\dot{K}_i(t) = \eta_i \, Z_i(t)^{1-\phi} \, S_i(t)^{\beta_i}, \qquad S_i(t) = \sum_j \phi_{ij}^{-\varsigma} K_j(t).$$

This is Proposition 2 in the main text:

$$F_i(Q,t) = \exp(-K_i(t) Q^{-\varsigma}), \qquad \dot{K}_i(t) = \eta_i Z_i(t)^{1-\phi} S_i(t)^{\beta_i}.$$

(Up to notational constants absorbed into η_i.)

Probability a new idea replaces the incumbent (explicit)

Fix a variety with current frontier Q. A new innovation attempt generates $Q_{\text{new}} = Q_n Q_o^{\beta_i}$.
Replacement event:

$$\{\text{replace}\} = \{Q_{\text{new}} > Q\}.$$

Conditional on $Q_o = q_o$, we have:

$$\mathbb{P}(\text{replace} \mid q_o) = \mathbb{P}\left(Q_n > \frac{Q}{q_o^{\beta_i}}\right) = \tilde{\eta}_i \, Q^{-\varsigma} \, q_o^{\beta_i \varsigma}.$$

Integrate over q_o:

$$\mathbb{P}(\text{replace}) = \tilde{\eta}_i Q^{-\varsigma} \mathbb{E}[q_o^{\beta_i \varsigma}] = \tilde{\eta}_i Q^{-\varsigma} \Gamma(1 - \beta_i) S_i(t)^{\beta_i}.$$

So: replacement hazard is proportional to $Q^{-\varsigma}$ (harder to beat high frontiers).

Patenting cutoff: map to a simple probability

Patenting rule: an innovation is patented only if its step size exceeds $\Psi_i(t)$:

$$\frac{Q_{\text{new}}}{Q} \geq \Psi_i(t).$$

Given Pareto tails, this produces a scaling:

$$P\dot{A}T_i(t) = \Psi_i(t)^{-\varsigma} \frac{\dot{K}_i(t)}{K_i(t)}.$$

Interpretation: $\Psi_i(t)$ shifts the observed patent rate without changing real innovation. (This is Lemma 1's patent equation.)

Citation probabilities: gravity form from the max problem

A patent in destination sector j cites the source sector i whose applicability-adjusted draw is maximal:

$$i = \arg\max_k \left\{ \frac{\tilde{Q}_{ok}}{\phi_{jk}} \right\}.$$

Compute:

$$\pi_{i|j}(t) \equiv \mathbb{P}\left(\frac{\tilde{Q}_{oi}}{\phi_{ji}} \geq \max_{k \neq i} \frac{\tilde{Q}_{ok}}{\phi_{jk}} \right).$$

With Fréchet draws, the standard EK result:

$$\pi_{i|j}(t) = \frac{\phi_{ji}^{-\varsigma} K_i(t)}{\sum_k \phi_{jk}^{-\varsigma} K_k(t)}.$$

This is the citation share equation in Lemma 1.

R&D arbitrage (free entry): where Z_i comes from

Let $V_i(t)$ be the expected value of owning a frontier variety in sector i. Free entry into R&D implies expected marginal cost equals expected marginal benefit:

$$\eta_i Z_i(t)^{-\phi} S_i(t)^{\beta_i} V_i(t) = 1.$$

Rearrange:

$$Z_i(t) = \left(\eta_i S_i(t)^{\beta_i} V_i(t)\right)^{1/\phi}.$$

Interpretation:

- **Push:** $\eta_i S_i^{\beta_i}$ (innovation productivity and spillovers)
- **Pull:** $V_i(t)$ (market size / profits path)

Reminder: HJB equations in continuous-time growth (general)*

In continuous time, dynamic optimization problems are characterized by a **Hamilton–Jacobi–Bellman (HJB) equation**.
Let $V(x,t)$ be the value function for a state $x(t)$ evolving as:

$$\dot{x}(t) = f(x(t), u(t), t).$$

Then $V(x,t)$ satisfies:

$$rV(x,t) = \max_u \left\{ \pi(x,u,t) + \partial_x V(x,t) \cdot f(x,u,t) + \partial_t V(x,t) \right\}.$$

Interpretation:

- left-hand side: opportunity cost of holding the asset
- right-hand side: flow payoff + capital gains

Where the HJB comes from (derivation sketch)*

Start from the Bellman equation over a small interval Δt:

$$V(x,t) = \max_u \left\{ \pi(x,u,t)\Delta t + e^{-r\Delta t} V\big(x + \dot{x}\,\Delta t,\; t + \Delta t\big) \right\}.$$

First-order expansion:

$$V(x + \dot{x}\,\Delta t, t + \Delta t) \approx V(x,t) + \partial_x V \cdot \dot{x}\,\Delta t + \partial_t V\,\Delta t.$$

Subtract $V(x,t)$ from both sides, divide by Δt, and let $\Delta t \to 0$:

$$rV(x,t) = \max_u \Big\{ \pi(x,u,t) + \partial_x V \cdot f(x,u,t) + \partial_t V(x,t) \Big\}.$$

Key point: HJB is just the continuous-time envelope condition.

HJB with Poisson arrival and creative destruction

In growth models with innovation, the state can change *discretely* via Poisson events.
Suppose:

- flow profit: $\pi(t)$
- discount rate: r
- innovation arrival with intensity $\lambda(t)$
- replacement probability upon arrival

Then the Bellman equation over Δt becomes:

$$V(t) = \pi(t)\Delta t + e^{-r\Delta t}\Big[(1-\lambda(t)\Delta t)V(t+\Delta t) + \lambda(t)\Delta t\, V^{\text{new}}(t+\Delta t)\Big].$$

Rearranging and taking $\Delta t \to 0$ yields:

$$rV(t) - \dot{V}(t) = \pi(t) - \lambda(t)\big(V(t) - V^{\text{new}}(t)\big).$$

Economic meaning:

- $\pi(t)$: flow monopoly profits
- $\lambda(t)$ term: expected capital loss from creative destruction

This is exactly the structure used in the paper.

From the general HJB to a per-idea HJB

Start from the standard HJB with Poisson creative destruction:

$$rV_i(t) - \dot{V}_i(t) = \pi_i(t) - \lambda_i(t)\big(V_i(t) - V_i^{\text{new}}(t)\big).$$

Key modeling choice in this paper:

- $V_i(t)$ is defined as the *value per frontier idea*
- Frontier ideas are replaced one-for-one by new ideas
- All frontier ideas are ex ante symmetric

As a result, creative destruction does not destroy value at the sector level: it reallocates value across ideas.

Why the hazard term disappears

Let $K_i(t)$ denote the mass of frontier ideas in sector i.

- Total sector value: $K_i(t)\,V_i(t)$
- Total sector profits: $\Pi_i(t)$
- Profit per frontier idea: $\pi_i(t) = \Pi_i(t)/K_i(t)$

Innovation replaces old ideas with new ones:

- individual ideas lose value,
- but the mass of ideas $K_i(t)$ grows over time.

Tracking value *per idea* absorbs creative destruction into $K_i(t)$. The HJB therefore simplifies to:

$$r(t)V_i(t) - \dot{V}_i(t) = \frac{\Pi_i(t)}{K_i(t)}.$$

Interpretation: faster growth of $K_i(t)$ shortens effective monopoly duration.

Value function dynamics: HJB-style ODE

Value per frontier idea evolves as:

$$r(t)V_i(t) - \dot{V}_i(t) = \frac{\Pi_i(t)}{K_i(t)}.$$

Interpretation of the RHS:

- $\Pi_i(t)$ is total sector profits (a revenue share times expenditure).
- Divide by $K_i(t)$ because K_i is the measure of frontier draws ("mass" of ideas).
- Faster innovation (higher K_i growth) shortens expected monopoly duration (competition effect).

Step 1: Define the object and the survival probability

Fixed varieties, evolving frontier distribution. Varieties are $v \in [0,1]$ (fixed). In sector i, the *frontier productivity* across varieties has Fréchet CDF

$$F_i(Q,t) = \Pr(Q_{iv}(t) \leq Q) = \exp\big(-K_i(t)\,Q^{-\theta}\big),$$

where $K_i(t)$ is the *Fréchet scale* ("knowledge stock") and $\theta > 0$ is the shape/dispersion.

Key max-stability implication. The frontier at date $s > t$ is the max of:

- the incumbent frontier draw at t, equal to Q;
- all new candidate draws arriving between t and s.

By Fréchet max-stability, the max of new draws over $(t,s]$ is Fréchet with scale $K_i(s) - K_i(t)$, hence:

$$\Pr(\text{no new draw exceeds } Q \text{ between } t \text{ and } s \mid Q) = \exp\big(-(K_i(s) - K_i(t))Q^{-\theta}\big).$$

This is the **survival probability** of an idea with quality Q from t to s.

Step 2: Value of an incumbent with quality Q and aggregation over Q

Let $\Pi_i(s)$ be **total sector profits** at time s. Since the frontier distribution is indexed by $K_i(s)$, the model's accounting implies the **flow profit per frontier draw** is

$$\pi_i(s) \equiv \frac{\Pi_i(s)}{K_i(s)}.$$

(See the appendix definition of V_i and the step converting profits into $\Pi_i(s)/K_i(s)$.)

Value conditional on frontier quality Q at time t:

$$\widetilde{V}_i(Q,t) = \int_t^\infty \pi_i(s) \underbrace{\exp\left(-\int_t^s r(\tau)\, d\tau \right)}_{\text{discount}} \underbrace{\exp\left(-(K_i(s) - K_i(t))Q^{-\theta} \right)}_{\text{survival}} ds.$$

Value per frontier idea (average over frontier draws at t):

$$V_i(t) \equiv \mathbb{E}_{Q\sim F_i(\cdot,t)}\left[\widetilde{V}_i(Q,t)\right] = \int_t^\infty \pi_i(s)\, e^{-\int_t^s r(\tau)\, d\tau} \underbrace{\mathbb{E}_{Q\sim F_i(\cdot,t)}\left[e^{-(K_i(s)-K_i(t))Q^{-\theta}}\right]}_{\star} ds.$$

Step 3: Compute the expectation $(\star)$ explicitly

We compute

$$(\star) = \mathbb{E}_{Q\sim F_i(\cdot,t)}\left[e^{-(K_i(s)-K_i(t))Q^{-\theta}}\right] = \int_0^\infty e^{-(K_i(s)-K_i(t))Q^{-\theta}}\, dF_i(Q,t).$$

Using $F_i(Q,t) = \exp(-K_i(t)Q^{-\theta})$, we have

$$dF_i(Q,t) = d\Big(\exp(-K_i(t)Q^{-\theta})\Big)\,.$$

Now apply the change of variables $y \equiv Q^{-\theta}$. Then $F_i(Q,t) = \exp(-K_i(t)y)$ and (as a Stieltjes integral)

$$\int_0^\infty e^{-(K_i(s)-K_i(t))Q^{-\theta}}\, dF_i(Q,t) = \int_{y=0}^\infty e^{-(K_i(s)-K_i(t))y}\, d\Big(e^{-K_i(t)y}\Big)\,.$$

But $d(e^{-K_i(t)y}) = -K_i(t)e^{-K_i(t)y}\, dy$, hence

$$(\star) = \int_0^\infty e^{-(K_i(s)-K_i(t))y}\, K_i(t)e^{-K_i(t)y}\, dy = K_i(t)\int_0^\infty e^{-K_i(s)y}\, dy = \frac{K_i(t)}{K_i(s)}.$$

This is the crucial "competition effect" term: faster growth in K_i lowers survival.

Step 4: Close form for $V_i(t)$ and the HJB-style ODE

Plugging $(\star) = K_i(t)/K_i(s)$ into the expression for $V_i(t)$:

$$V_i(t) = \int_t^\infty \pi_i(s)\, e^{-\int_t^s r(\tau)\,d\tau}\, \frac{K_i(t)}{K_i(s)}\, ds = \int_t^\infty \frac{\Pi_i(s)}{K_i(s)}\, e^{-\int_t^s r(\tau)\,d\tau}\, ds.$$

This is precisely the appendix definition (eq. (77) there) and it *implies* eq. (23).

Differentiate to get the ODE. Let $D(t,s) \equiv e^{-\int_t^s r(\tau)\,d\tau}$. Then

$$V_i(t) = \int_t^\infty \frac{\Pi_i(s)}{K_i(s)} D(t,s)\, ds.$$

Differentiate using Leibniz rule:

$$\dot{V}_i(t) = -\frac{\Pi_i(t)}{K_i(t)} + \int_t^\infty \frac{\Pi_i(s)}{K_i(s)} \frac{\partial}{\partial t} D(t,s)\, ds.$$

Since $\frac{\partial}{\partial t} D(t,s) = r(t) D(t,s)$, we obtain

$$\dot{V}_i(t) = -\frac{\Pi_i(t)}{K_i(t)} + r(t) V_i(t) \quad \Longrightarrow \quad r(t) V_i(t) - \dot{V}_i(t) = \frac{\Pi_i(t)}{K_i(t)}.$$

This shows the result in the slide “Value function dynamics: HJB-style ODE.”

Profit decomposition (price, competition, income effects)

Using demand + pricing, the paper shows relative profits per idea satisfy:

$$\frac{\Pi_i(t)/K_i(t)}{\Pi_j(t)/K_j(t)} = \underbrace{\frac{K_j(t)}{K_i(t)}}_{\text{competition}} \cdot \underbrace{\left(\frac{K_j(t)}{K_i(t)}\right)^{1-\sigma}}_{\text{price}} \cdot \underbrace{C(t)^{(1-\sigma)(\varepsilon_i-\varepsilon_j)}}_{\text{income}}.$$

Three effects:

- competition: faster innovation lowers duration
- price: higher K lowers price index (depending on σ)
- income: nonhomotheticity drives shifting demand

Sectoral price index: definition

Variety price under Bertrand competition with fringe vintage:

$$P_{iv}(t) = \frac{\chi}{Q_{iv}(t)}.$$

Sectoral output uses a log aggregator:

$$Y_i(t) = \exp\left(\int_0^1 \log X_{iv}(t)\, dv\right), \quad X_{iv}(t) = Q_{iv}(t)L_{iv}(t).$$

Hence the sectoral price index is the geometric mean:

$$\log P_i(t) = \int_0^1 \log P_{iv}(t)\, dv = \log \chi - \int_0^1 \log Q_{iv}(t)\, dv.$$

With a continuum of varieties and i.i.d. frontier draws:

$$\log P_i(t) = \log \chi - \mathbb{E}[\log Q_i(t)].$$

Log-moment of the Fréchet distribution

Suppose frontier productivities satisfy:

$$F_i(Q,t) = \exp\left(-K_i(t)Q^{-\varsigma}\right).$$

Define the transformation:

$$U \equiv K_i(t)Q^{-\varsigma}.$$

Then:

$$\Pr(U > u) = \Pr\Big(Q < (K_i/u)^{1/\varsigma}\Big) = F_i\Big((K_i/u)^{1/\varsigma}, t\Big) = e^{-u},$$

so:

$$U \sim \mathrm{Exp}(1).$$

Hence:

$$\log Q = \frac{1}{\varsigma}\left(\log K_i(t) - \log U\right), \qquad \mathbb{E}[\log U] = -\gamma_E,$$

where γ_E is the Euler–Mascheroni constant.

Closed-form sectoral price index

Taking expectations:

$$\mathbb{E}[\log Q_i(t)] = \frac{1}{\varsigma}\left(\log K_i(t) + \gamma_E\right).$$

Substitute into the price index:

$$\log P_i(t) = \log \chi - \frac{1}{\varsigma}\left(\log K_i(t) + \gamma_E\right).$$

Exponentiating:

$$P_i(t) = \chi \exp\left(-\frac{\gamma_E}{\varsigma}\right) K_i(t)^{-1/\varsigma}.$$

Key implication: all time-variation in prices is driven by $K_i(t)$.

Price inflation and knowledge growth

Let $p_i(t) \equiv \log P_i(t)$ and $k_i(t) \equiv \log K_i(t)$.
From the closed form:

$$p_i(t) = \log \chi - \frac{\gamma_E}{\varsigma} - \frac{1}{\varsigma} k_i(t).$$

Differentiate:

$$\dot{p}_i(t) = -\frac{1}{\varsigma}\dot{k}_i(t) = -\frac{1}{\varsigma}\frac{\dot{K}_i(t)}{K_i(t)}.$$

Interpretation: faster idea accumulation $\Rightarrow$ faster relative price declines.

Model objects and observables

Model objects (sector *i*):

- $K_i(t)$: Fréchet scale of frontier productivities (knowledge stock)
- $\dot{K}_i(t)/K_i(t)$: growth rate of frontier knowledge
- ϕ_{ij}: applicability of knowledge from sector j to sector i
- β_i: elasticity of innovation productivity to external knowledge
- $\Psi_i(t)$: patenting cutoff (minimum quality improvement to patent)

Observed data:

- Patent counts by sector and year
- Citation shares across sectors
- Sectoral R&D expenditure
- Sectoral output and expenditure shares

Goal: map patents and citations into $K_i(t)$, β_i, ϕ_{ij}, and $\Psi_i(t)$.

Patents as selected frontier innovations

In the model, innovation arrivals are Poisson. Each arrival generates a quality improvement:

$$\Delta \equiv \frac{Q_{\text{new}}}{Q_{\text{old}}}.$$

Assumption (tail behavior): Quality improvements have Pareto tails:

$$\Pr(\Delta \geq x) = x^{-\theta}, \qquad x \geq 1.$$

Patenting rule: An innovation in sector i is patented if:

$$\Delta \geq \Psi_i(t),$$

where $\Psi_i(t) \geq 1$ is a sector– and time–specific patenting cutoff.
Probability an innovation is patented:

$$\Pr(\text{patent}_i \mid t) = \Pr(\Delta \geq \Psi_i(t)) = \Psi_i(t)^{-\theta}.$$

Key implication: Patents are a *selected sample* of frontier innovations.

Recovering innovation rates from patent data

Let:

- $\dot{K}_i(t)$ = flow of frontier innovations (model object),
- $\mathrm{PAT}_i(t)$ = observed patent counts (data).

Because only a fraction $\Psi_i(t)^{-\theta}$ of innovations are patented:

$$\mathrm{PAT}_i(t) = \Psi_i(t)^{-\theta}\,\dot{K}_i(t).$$

Equivalently,

$$\frac{\mathrm{PAT}_i(t)}{\dot{K}_i(t)} = \Psi_i(t)^{-\theta}.$$

Interpretation:

- Patent counts do *not* identify $\dot{K}_i(t)$ directly.
- They identify $\dot{K}_i(t)$ *up to a selection wedge* $\Psi_i(t)^{-\theta}$.
- Changes in patenting standards or behavior are absorbed by $\Psi_i(t)$.

Empirical strategy: Allow $\Psi_i(t)$ to vary over time and sector, and use the model structure to separately discipline $\dot{K}_i(t)$.

How the knowledge stock $K_i(t)$ is computed

Step 1: What patents measure (flows). Patent counts in sector i at time t satisfy:

$$\mathrm{PAT}_i(t) = \Psi_i(t)^{-\theta}\,\dot{K}_i(t),$$

- $\dot{K}_i(t)$ = flow of frontier innovations (model object),
- $\Psi_i(t)$ = patenting cutoff (selection into patents),
- θ = Pareto tail parameter.

Thus, patents identify *innovation flows up to a selection wedge*.

Step 2: Back out innovation flows. Given $\Psi_i(t)$,

$$\dot{K}_i(t) = \Psi_i(t)^{\theta}\,\mathrm{PAT}_i(t).$$

Step 3: Accumulate flows into a stock.
The Fréchet scale (knowledge stock) is constructed as:

$$K_i(t) = K_i(0) + \int_0^t \dot{K}_i(s)\,ds = K_i(0) + \int_0^t \Psi_i(s)^{\theta}\,\mathrm{PAT}_i(s)\,ds.$$

Key point: $K_i(t)$ is a *latent stock* inferred by accumulating patent-adjusted innovation flows. Its level is identified up to a normalization, but ratios and growth rates are pinned down.

Citation probabilities and applicability parameters

In the model, a new innovation in sector j builds on the best available idea:

$$i = \arg\max_k \left\{ \frac{\tilde{Q}_{ok}}{\phi_{jk}} \right\}.$$

With Fréchet frontier draws, the probability that sector j cites sector i is:

$$\pi_{ij}(t) = \Pr(i \mid j) = \frac{\phi_{ij}^{-\theta} K_i(t)}{\sum_k \phi_{kj}^{-\theta} K_k(t)}.$$

Key result: Citation shares satisfy a *gravity equation* in knowledge stocks and applicability costs.

Estimating applicability ϕ_{ij} from citations

Taking logs relative to within-sector citations:

$$\log \pi_{ij}(t) - \log \pi_{jj}(t) = -\theta \log \phi_{ij} + \log K_i(t) - \log K_j(t).$$

Empirical implementation:

- $\pi_{ij}(t)$ observed from citation matrices
- $K_i(t)$ recovered from patent-adjusted innovation flows
- θ taken from the literature or estimated

This identifies ϕ_{ij} up to normalization $\phi_{jj} = 1$.
Interpretation:

- lower $\phi_{ij} \Rightarrow$ ideas from i are more applicable to j
- manufacturing typically has low ϕ_{ij} across many sectors

Recovering spillover elasticities β_i

The innovation technology in sector i is:

$$\dot{K}_i(t) = \eta_i Z_i(t)^{1-\alpha} \left(\sum_j \phi_{ij}^{-\theta} K_j(t) \right)^{\beta_i} .$$

Define the spillover index:

$$S_i(t) \equiv \sum_j \phi_{ij}^{-\theta} K_j(t).$$

Taking logs and differences:

$$\Delta \log \dot{K}_i(t) = (1-\alpha)\Delta \log Z_i(t) + \beta_i \Delta \log S_i(t) + \Delta \log \eta_i.$$

Given:

- $\dot{K}_i(t)$ from patent-adjusted innovation rates
- $Z_i(t)$ from R&D data
- $S_i(t)$ from citation-weighted knowledge stocks

we identify the spillover elasticity β_i.

Mapping data to model objects (corrected)

Model object	Identified from data
$\mathrm{PAT}_i(t)$	Observed patent counts (data)
$\dot{K}_i(t)$	Patent counts adjusted for selection $\Psi_i(t)$
$K_i(t)$	Time integral of $\dot{K}_i(t)$ (normalized)
$\Psi_i(t)$	Patenting intensity conditional on R&D and output
ϕ_{ij}	Cross-sector citation shares $\pi_{ij}(t)$
β_i	Response of $\dot{K}_i(t)$ to spillover index $S_i(t)$

Interpretation:

- Patents identify *flows* of frontier innovation, not knowledge stocks.
- Knowledge stocks $K_i(t)$ are inferred by accumulating those flows.
- Citations depend on relative $K_i(t)$, so normalization does not matter.

State variables, controls, and equilibrium system

State variables (knowledge stocks):

$$k_i(t) \equiv \log K_i(t), \qquad i = 1, \ldots, I.$$

Controls:

$$c(t) \equiv \log C(t), \qquad \lambda_i(t) \equiv \frac{Z_i(t)}{Z(t)}, \quad \sum_i \lambda_i(t) = 1.$$

Equilibrium dynamics can be written compactly as:

$$\dot{x}(t) = f\big(x(t), u(t), t; \theta\big),$$

where:

$$x(t) = (k_1, \ldots, k_I, c), \qquad u(t) = (\lambda_1, \ldots, \lambda_{I-1}).$$

Why collocation instead of shooting

Main challenges:

- High-dimensional controls $\lambda_i(t)$
- Non-autonomous dynamics (expenditure shares shift over time)
- Stable manifold selection is non-trivial

Collocation strategy:

- Approximate unknown paths with smooth splines
- Enforce equilibrium conditions pointwise on a grid
- Impose initial and terminal (CGP) conditions directly

This avoids instability and sensitivity typical of shooting methods.

Spline approximation of paths

Choose a horizon $[0, T]$ and knots:

$$0 = \tau_0 < \tau_1 < \cdots < \tau_M = T.$$

Approximate paths using B-splines:

$$k_i(t) \approx \sum_{m=1}^{M_k} a_{im} B_m(t), \qquad c(t) \approx \sum_{m=1}^{M_c} b_m B_m(t).$$

Enforce simplex constraints on R&D shares via softmax:

$$\lambda_i(t) = \frac{\exp\left(\sum_m d_{im} B_m(t)\right)}{\sum_j \exp\left(\sum_m d_{jm} B_m(t)\right)}.$$

Derivatives are analytical:

$$\dot{k}_i(t) = \sum_m a_{im} \dot{B}_m(t), \quad \dot{c}(t) = \sum_m b_m \dot{B}_m(t).$$

Collocation residuals

Model-implied dynamics:

$$\dot{k}_i(t) = g_i(k(t), c(t), \lambda(t), t; \theta), \qquad \dot{c}(t) = g_c(\cdot).$$

Define residuals at time t:

$$R_{k,i}(t) = \dot{k}_i^{\text{approx}}(t) - g_i(\cdot), \qquad R_c(t) = \dot{c}^{\text{approx}}(t) - g_c(\cdot).$$

Only $I - 1$ share equations are needed, since $\sum_i \lambda_i(t) = 1$ is enforced by construction.

Objective function: nonlinear least squares

Choose collocation nodes $\{t_n\}_{n=1}^N$ (e.g. Gauss–Lobatto).
Solve:

$$\min_{a,b,d} \sum_{n=1}^{N} w_n \left[\sum_{i=1}^{I} R_{k,i}(t_n)^2 + R_c(t_n)^2 + \sum_{i=1}^{I-1} R_{\lambda,i}(t_n)^2 \right].$$

This is a standard nonlinear least-squares problem over spline coefficients.

Boundary and terminal conditions

Initial conditions (from calibration/data/searched over really in the code):

$$k_i(0) = k_{i0}.$$

Terminal anchoring to the constant-growth path (CGP):

$$\dot{k}_i(T) \approx \dot{k}_i^{\text{CGP}}, \quad \dot{c}(T) \approx g^*, \quad \lambda_i(T) \approx \lambda_i^{\text{CGP}}.$$

Impose via penalty terms:

$$\Omega_T = \omega_k \sum_i (\dot{k}_i(T) - \dot{k}_i^{\text{CGP}})^2 + \omega_c(\dot{c}(T) - g^*)^2 + \omega_\lambda \sum_i (\lambda_i(T) - \lambda_i^{\text{CGP}})^2.$$

Choice of knots and collocation nodes

Practical recommendations:

- Use a **non-uniform knot grid**:
  - dense near $t = 0$ (fast transitions)
  - sparse near T (CGP convergence)
- Typical choice: $M = 15$–25 knots per state
- Use Gauss–Lobatto nodes within each interval

This balances accuracy and computational cost.

Normalization for numerical stability

Normalize by CGP trends:

$$\tilde{k}_i(t) = k_i(t) - \dot{k}_i^{\text{CGP}} t, \qquad \tilde{c}(t) = c(t) - g^* t.$$

Then terminal conditions become:

$$\dot{\tilde{k}}_i(T) \approx 0, \qquad \dot{\tilde{c}}(T) \approx 0.$$

This dramatically improves conditioning and convergence.

Jacobian computation and solvers

Implementation tips:

- Use automatic differentiation if available
- Otherwise, exploit analytical derivatives of splines
- Feed Jacobian to a trust-region or Levenberg–Marquardt solver

Collocation works best with:

- accurate Jacobians
- tight but not excessive terminal penalties

Diagnostics and convergence checks

Always report:

- max residual norm: $\max_n \|R(t_n)\|$
- drift at terminal time: $\|\dot{x}(T) - \dot{x}^{\mathrm{CGP}}\|$
- robustness to:
  - more knots
  - longer horizon T
  - alternative initial guesses

A good solution is stable to all three.

Takeaways

- Fréchet structure gives closed-form prices and clean dynamics
- Transitional dynamics are smooth but high-dimensional
- Collocation turns equilibrium into a transparent residual-minimization problem
- Every numerical step has an economic interpretation