Markup Dynamics and Firm Entry

Eero Mäkynen [University of Turku] Philip Schnattinger [Bank of England]

RESEARCH QUESTION & KEY FINDINGS

How do markups evolve over the firm life cycle, and what does this imply for efficient firm entry?

We show that:

1. **Markup dynamics are mostly ex-post.** Using rich UK firm-level data, we decompose log markups into an ex-ante component fixed at entry and ex-post shocks realised over the life cycle. Ex-post shocks account for most cross-sectional dispersion and higher moments of markups.
2. **Markup dynamics matter for optimal entry.** We solve a heterogeneous-firm model with firm-specific markups, disciplined to match markup covariances and moments by firm age. In this environment, the equilibrium number of firms can differ from the planner's allocation, and the entry wedge depends on whether markup dispersion is mainly predetermined at entry or instead accumulates through life-cycle shocks.
3. **Result.** Optimal entry subsidies can raise output and welfare, but the size of the gain varies systematically with the balance between ex-ante and ex-post markup heterogeneity and can become a loss depending on the firm-age-markup structure of the economy.

Data & Measurement

Data. Firm-level panel data for the United Kingdom, 1994–2024, covering the private non-financial sector, with ≈ 1.2 million firm-year observations and ≈ 350 thousand distinct firms.

Markup measure. We construct several markup measures; our baseline proxy is revenue/expenditure based:

$$\mu_{it} = \log\left(\frac{\text{revenue}_{it}}{\text{flexible input expenditure}_{it}}\right).$$

To focus on within-market dispersion, we remove industry and calendar-year fixed effects and use the residual as our log markup. Firm age is measured as years since first appearance in the data.

Life-cycle patterns. Over the first 15 years of firm age, $\mathrm{Var}(\mu)$ declines from about 1.1 to 0.9, skewness rises from roughly 0 to 1, and kurtosis stays close to 8. These non-Gaussian patterns motivate both our decomposition and the quantitative model.

Empirical Strategy

Ex-ante / ex-post decomposition. In the spirit of Sterk, Sedláček, and Pugsley (2021), we decompose log markups into ex-ante (*determined at birth*) and ex-post (*realised over the life cycle*) components:

$$\mu_t = \mu_{d,t} + \mu_{p,t} + \mu_{\varepsilon,t},$$

where $\mu_{d,t}$ is an ex-ante deterministic path from an initial draw to a firm-specific endpoint, $\mu_{p,t}$ is a persistent ex-post AR(1) component, and $\mu_{\varepsilon,t}$ is i.i.d. noise. Parameters are chosen to match the age profile and autocovariance surface of log markups, yielding a variance decomposition into ex-ante and ex-post parts.

Higher moments. We then add parameters that affect only the third and fourth cumulants, identified from the age profiles of skewness and kurtosis while keeping the covariance structure fixed. This lets us ask whether asymmetry and fat tails in markups are mainly driven by ex-ante heterogeneity or by ex-post life-cycle shocks, without altering the standard variance decomposition.

Component Estimates & Decomposition of Central Moments

A Gaussian shocks mixed with Poisson–exponential jumps in the initial ex-ante draw and in the persistent component fits the variance and higher moments of markups. Jump intensities are low, so most firms evolve smoothly, but occasional events move firms sharply into or out of the right tail of the markup distribution.

Table 1. Empirical results

Baseline covariance structure		Extended shock structure for higher moments	
Component	Estimate	Component	Estimate
Ex-post persistence (ρ)	0.789	Ex-post jump sizes (θ_p)	1.714
Ex-post shock variance (V_p)	0.231	Ex-post jump intensity (λ_p)	0.013
Ex-ante persistence (ϕ)	0.801	Ex-post residual shock (σ_p^2)	0.155
Ex-ante shock variance (V_{d0})	0.850	Ex-ante jump sizes (θ_{d0})	1.314
IID noise (σ_ε^2)	0.513	Ex-ante jump intensity (λ_{d0})	0.089
Endpoint heterogeneity (σ_γ^2)	0.036	Ex-ante residual shock (σ_{d0}^2)	0.542

Endpoint heterogeneity and i.i.d. noise are restricted to be Gaussian; these are the only parametric assumptions needed for the decomposition itself. The Poisson–exponential parameters are estimated only as one example of a process consistent with the implied higher moments, abstracting from firm selection; selection is treated in the full quantitative model.

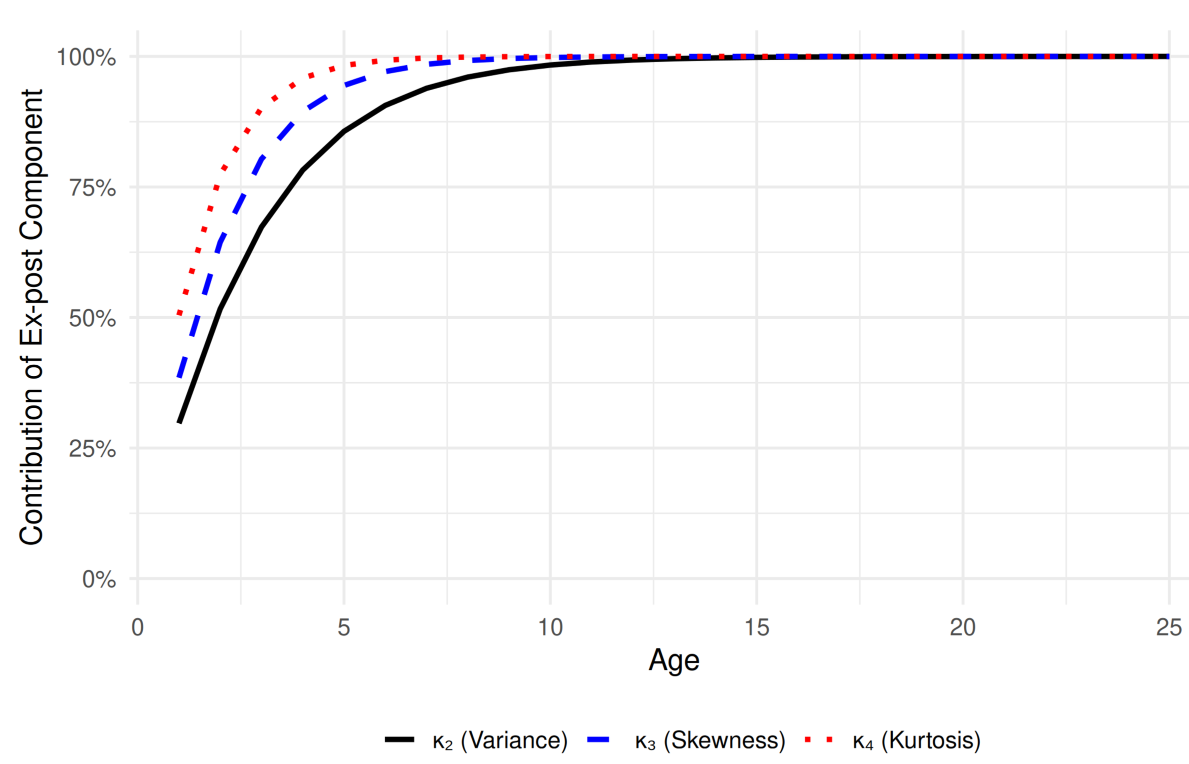


Figure 1. The relative contribution of ex-post components to cross-sectional moments of markups over firm age.

Key Model Elements

We use a **Hopenhayn-style** dynamic industry model with heterogeneous firms and **firm-specific markups**. **Varieties face HSA demand** (Baqaee, Farhi, & Sangani, 2023; Matsuyama, 2025), which nests CES and links relative prices φ to demand elasticities $\epsilon(\varphi)$ and markups via $\mu(\varphi) = \epsilon(\varphi)/(\epsilon(\varphi) - 1)$.

Each firm has **idiosyncratic productivity** z that determines marginal costs and hence its optimal relative price and markup,

$$z \mapsto \varphi(z) \mapsto \mu(\varphi(z)),$$

so productivity dispersion translates into markup dispersion. The productivity process follows a Markov structure calibrated to reproduce the empirical markup dynamics.

Entry and exit follow Hopenhayn and Rogerson (1993): potential entrants decide whether to pay an entry cost, and incumbents choose optimally when to exit. These decisions generate a stationary measure of productivities Ψ over z, and the cross-sectional markup distribution is the pushforward of Ψ under $z \mapsto \mu(\varphi(z))$. This setup allows us to compare equilibrium and planner allocations under different sources of markup dispersion.

Productivity- Markup- Process specification. Intuitively, $z_{d,t}$ captures the ex-ante life-cycle profile, $z_{p,t}$ captures persistent ex-post shocks, and the jump component X captures investment in markups which generates rare large movements that produce skewness and fat tails in productivity and hence markups.

$$z_t = z_{d,t} + z_{p,t} + z_{\varepsilon,t}, \qquad z_{\varepsilon,t} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$$

$$z_t = z_{d,t} + z_{p,t} + z_{\varepsilon,t}, \qquad z_{p,t+1} = \rho\, z_{p,t} + u_t, \qquad z_{d,t+1} = \phi\, z_{d,t} + \theta_0,$$
$$z_{\varepsilon,t} \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad u_t, z_{d,0} \sim X, \quad X = \epsilon + \left(\sum_{i=1}^{N} Y_i - \mathbb{E}\Big[\sum_{i=1}^{N} Y_i\Big]\right), \quad \epsilon \sim \mathcal{N}(0, \sigma^2), \quad N \sim \text{Pois}(\lambda),\ Y_i \sim \text{Exp}(\tfrac{1}{\theta}).$$

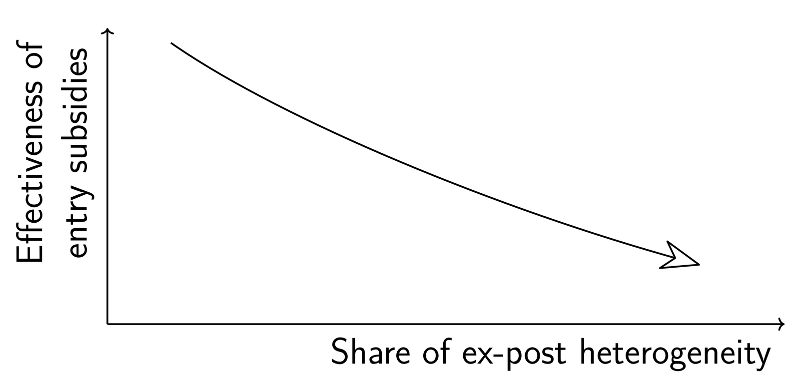
Markup Dynamics and Optimal Entry

Intuition. With stochastic productivity, incumbents value the option of waiting for better future draws, which keeps some low-productivity firms active. A planner internalises how entry and exit jointly shape the productivity and markup distribution, so the equilibrium number of firms need not coincide with the planner's choice. Under HSA demand, productivity dispersion implies that markups vary systematically over the firm life cycle, and the welfare effect of entry depends on whether most markup dispersion appears already at entry or builds up later.

Quantitative illustration. We compare economies with similar overall dispersion but different splits between ex-ante and ex-post heterogeneity. The table reports output gains from the optimal entry subsidy relative to the decentralised equilibrium.

Output gains from optimal entry subsidies

Heterogeneity case	Output gain ΔY (%)
More ex-ante heterogeneity	4.15
Baseline	3.95
More ex-post heterogeneity	3.15



Takeaways

- Holding overall dispersion roughly fixed, shifting heterogeneity from ex-post shocks to ex-ante entry draws **raises** the gains from optimally correcting the entry margin.
- When ex-ante heterogeneity is more prominent, the entry margin can be used to increase firm turnover and partially offset the dynamic externality created by incumbents' option value of waiting.
- Under HSA demand, this comes at the cost of greater markup dispersion, which dampens part of the potential gains, so optimal entry policy delivers only **modestly** higher output gains when heterogeneity is primarily ex-ante than when it is primarily ex-post.

More generally, the strength of the ex-ante / ex-post contrast depends on three elements: (i) the age structure of firms encoded in Ψ, (ii) departures from CES implied by HSA demand, and (iii) the tail behaviour of productivity shocks. Different combinations of these ingredients change how much of the inefficiency can be addressed at the entry margin, relative to instruments that act directly on incumbents.

Additional Results & Next Steps

In the paper, we also separate the role of markup heterogeneity from the standard CES variety channel by comparing economies with and without HSA demand, and by varying the balance between ex-ante and ex-post dispersion. This helps to isolate when entry wedges are driven mainly by dynamic selection and when they are amplified or dampened by markup dispersion.

In ongoing work we study alternative markup measures, sectoral differences in markup dynamics and entry wedges, and robustness of the decomposition to alternative shock specifications. We also explore how other policy instruments interact with markup dynamics, beyond simple entry subsidies.

Disclaimer & References

Baqaee, D. R., Farhi, E., & Sangani, K. (2023). The darwinian returns to scale. *The Review of Economic Studies*, *91*(3), 1373-1405.

Hopenhayn, H., & Rogerson, R. (1993). Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*, *101*(5), 915-938.

Matsuyama, K. (2025). Homothetic non-ces demand systems with applications to monopolistic competition. *Annual Review of Economics*, *17*, 261-292.

Sterk, V., Sedláček, P., & Pugsley, B. (2021). The nature of firm growth. *American Economic Review*, *111*(2), 547–79.