A stochastic SIR model for cyber contagion: application to granular growth of firms and insurance portfolio

Caroline Hillairet Olivier Lopez **Lionel Sopgoui**

CREST, Ensae Paris & Institut Louis Bachelier

7th European Actuarial Journal Conference 2026 - September 10, 2026

Cybersecurity: A Growing Concern

- Cyber risk ranks among top emerging threats (climate change, geopolitical instability).
- Ransomware attacks: explosive growth
  - 76.8% increase (2022–2023); 58% increase (2024–2025) $\rightarrow$ 7,515 publicly reported attacks in 2025.
  - \$1.1 billion paid to hackers in 2023 ([AXA Group, 2025]).
- Need to assess financial impact on firms and insurance portfolios.

- Firm size distributions are heavy-tailed (Zipf's law) [Gabaix, 1999].
- Firm growth-rate distributions are non-Gaussian ([Stanley et al., 1996, Axtell, 2001]).

Firm data: 2,929 firms in Île-de-France; annual revenues 2010–2022 (Pappers).

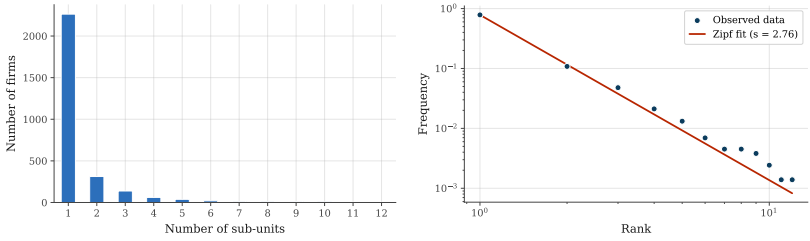

Figure: Firm size distribution (left) and Zipf fit (right).

- Firm size K_i estimated from revenue: $K_i = \lceil Z_{i,1}/\overline{Z}_1 \rceil$.
- Zipf's law fit: $\mathbb{P}(K_i = k) \propto k^{-(1+\alpha)}$ with $\alpha = 1.76$, $q = 0.784$.

- Larger firms are more frequently targeted [Kamiya et al., 2021, Baksy et al., 2025].
- Contagion dynamics depend on firm size and organizational structure.
- Environmental variability affects parameters [European Central Bank, 2025].

Cyber contagion data: LockBit ransomware attacks, May–July 2024 ([GuidePoint Security, 2025]).

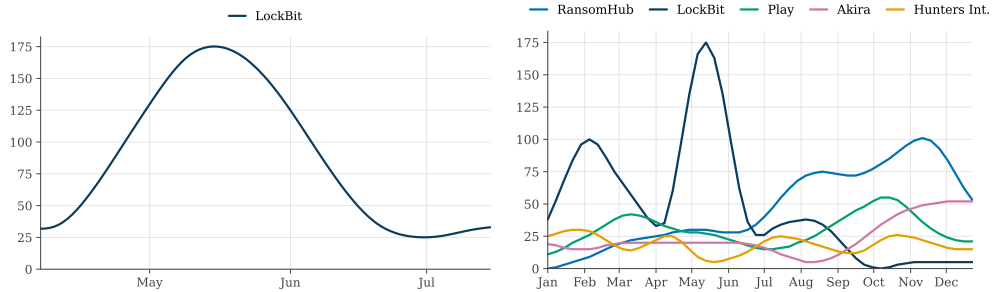

Figure: LockBit infections (left) and most impactful ransomware groups (right).

Goal: Build a model incorporating these features to quantify cyber risk.

1. **Granular model of firm growth**
   ([Wyart and Bouchaud, 2003, Moran et al., 2024])
   - Firms composed of independent subunits.
   - Revenue follows jump-diffusion with cyber-attack losses.
2. **Stochastic multigroup SIR model** for cyber contagion
   - Firms grouped by size $k = 1, \ldots, K$.
   - Stochastic parameters (Cox–Ingersoll–Ross processes) capture environmental variability.
3. **Coupling:** Attack arrival via Cox process (external) + Bernoulli (internal contagion).
4. **Risk measure:** Aggregate Exceedance Probability (AEP) and Systemic Tail Ratio for insurance portfolios.

The model

Firm revenue

$$Z_{i,t} = \sum_{j=1}^{K_i} z_{ij,t}, \quad \text{with } K_i \text{ the number of subunits (Zipf on } \{1, \ldots, K\}).$$

Without attack: Geometric Brownian motion $\dfrac{d\overline{z}_{ij,t}}{\overline{z}_{ij,t}} = \mu_{ij}dt + \sigma_{ij}dB_{ij,t}.$

With attack (arrival τ_{ij}, duration δ_{ij}, severity π_{ij}):

$$z_{ij,t} = \begin{cases} \overline{z}_{ij,t} & t < \tau_{ij} \\ (1-\pi_{ij})\overline{z}_{ij,t} & \tau_{ij} \le t < \tau_{ij} + \delta_{ij} \\ \overline{z}_{ij,t} & t \ge \tau_{ij} + \delta_{ij} \end{cases}, \qquad \tau_{ij} = \inf\{t \ge 0 \mid N_{ij,t} \ge 1\},$$

where $N_{ij,t}$ is a point process (defined later). Subunits are correlated within a firm, independent across firms.

Instantaneous claim for firm i:

$$c_{i,t} = \sum_{j=1}^{K_i} \pi_{ij} \mathbf{1}_{\{\tau_{ij} \le t < \tau_{ij} + \delta_{ij}\}} \overline{z}_{ij,t}, \qquad \mathfrak{C}_T = \sum_{i=1}^{H} \int_0^T c_{i,t} \, \mathrm{d}t.$$

Aggregate Exceedance Probability (AEP):

$$\mathrm{AEP}_T(x) = \mathbb{P}\left(\sum_{p=1}^{P} \mathfrak{C}_T^p > x \right),$$

where P is the number of independent cyber-events and $\mathfrak{C}_T^p$ are i.i.d. copies of $\mathfrak{C}_T$.
Systemic Tail Ratio:

$$\mathcal{R}_{\mathrm{sys}}(\alpha) := \frac{\mathrm{ES}_\alpha(\mathfrak{C}_T)}{\sum_{i=1}^{H} \mathrm{ES}_\alpha(\int_0^T c_{i,t} \, dt)}, \qquad 0 \le \alpha \le 1,$$

where ES_α is the Expected Shortfall at level α.

Literature Review: Modeling Cyber-Risk Contagion

- **Frequency-Severity Models**:
  [Eling and Loperfido, 2017, Farkas et al., 2021, Edwards et al., 2016].
- **Self-Exciting Point Processes**:
  [Bessy-Roland et al., 2021, Boumezoued et al., 2025].
- **Agent-Based & Network Models**:
  [Acemoglu et al., 2016, Fahrenwaldt et al., 2018].
- **Epidemiological (SIR) Models**: [Hillairet and Lopez, 2021, Hillairet et al., 2022, Murray, 1988, Langlois-Berthelot et al., 2024].

How Cyber-Infection Spreads into Subunits

- A firm of size j can be **susceptible**, **infected**, or **recovered**.
- An initial infection in a susceptible firm triggers **secondary infections** inside: each of the $j-1$ other subunits is infected with probability a_t (in-firm rate).
- Number of infected subunits is **binomial**: $b_{j,k} = \binom{j-1}{k-1} a^{k-1}(1-a)^{j-k},\ k = 1, \ldots, j.$
- The firm splits into a fully infected firm of size k and a susceptible firm of size $j-k$.
- **External contagion** between firms is driven by the force of infection; infected firms recover with rate $\gamma_{k,t}$.

Compartmental Model with Stochastic Parameters

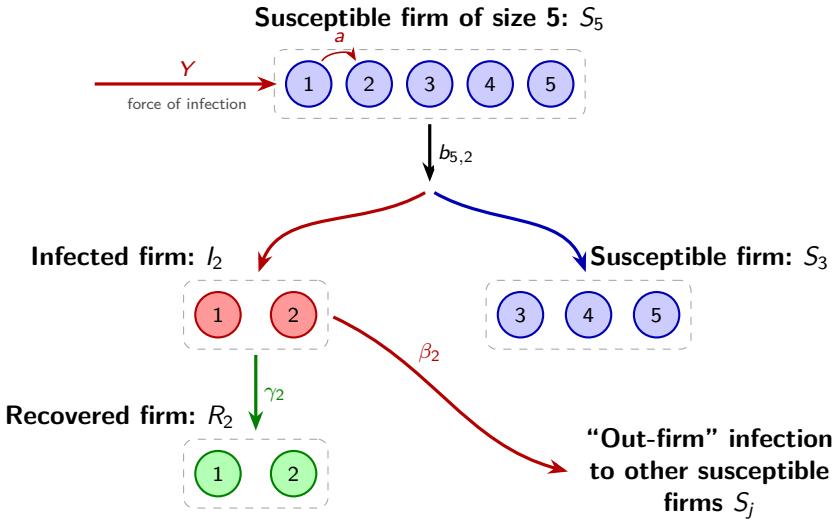

Figure: Infection and splitting of a firm of size 5: $S_5 \to I_2 + S_3$.

S_k, I_k, R_k: fractions of susceptible/infected/recovered firms of size k;
$N_0 = \sum_k k(S_k + I_k + R_k)$. Force of infection: $Y_t = \frac{1}{N_0} \sum_{k=1}^{K} \beta_{k,t}\, k\, I_{k,t}$.

$$\begin{cases} \frac{\mathrm{d}S_{k,t}}{\mathrm{d}t} = Y_t \left(-kS_{k,t} + \sum_{j=k+1}^{K} jS_{j,t}\, b_{j\,j-k,t} \right), \\ \frac{\mathrm{d}I_{k,t}}{\mathrm{d}t} = -\gamma_{k,t} I_{k,t} + Y_t \sum_{j=k}^{K} jS_{j,t}\, b_{jk,t} \\ \frac{\mathrm{d}R_{k,t}}{\mathrm{d}t} = \gamma_{k,t} I_{k,t}, \end{cases}$$

Environmental variability: β_k, γ_k, a follow CIR processes

$$d\tilde{\varphi}_t = \kappa_\varphi(\mu_\varphi - \tilde{\varphi}_t)dt + \Sigma_\varphi \sqrt{\tilde{\varphi}_t}\, dW_{\varphi,t}, \qquad \varphi = 1/(1 + e^{-\tilde{\varphi}}) \text{ for probabilities.}$$

Existence/uniqueness of a non-negative solution + reproduction number $\mathcal{R}_{k,t}$: see backup.

Definition of the Attack Arrival Process

$$N_{ij,t} := N^0_{ij,t} + \mathbf{1}_{\left\{N^0_{ij,\tau_i^-}=0\right\}} \mathbf{1}_{\{t \geq \tau_i\}} U_{ij},$$

- N^0_{ij}: Cox process with intensity Y_t (force of infection) – **external infections**.
- $\tau_i = \inf\{t > 0 : \sum_{k=1}^{K_i} N^0_{ik,t} \geq 1\}$: first external infection in firm i.
- $U_{ij} \sim \text{Bernoulli}(a_{\tau_i})$: **internal** infection of a still-susceptible sister subunit.
- Recovery duration: $\delta_{ij} = 1/\gamma_{i,\tau_{ij}}$.

External infection: Cox process

↓

First external infection in firm i at τ_i

↓

Other subunits: infected with probability a_{τ_i}

Numerical Analysis and Results

We minimize the mean squared error $J_2(\Theta) = \frac{1}{M}\sum_{m=1}^{M}\sum_{k=1}^{K}\sum_{u=0}^{T}|I_{k,t_u} - I_{k,t_u}^{m}|^2$, with $\Theta = \{\kappa, \Sigma, a_0, \gamma_{1,0}, \beta_{1,0}, h_\star\}$.

Parameter	Value	Description
$\gamma_{1,0}$	0.678	Recovery rate
$\beta_{1,0}$	0.547	Out-firm infection rate
a_0	0.347	In-firm attack rate
κ	0.447	Mean reversion speed
Σ	0.015	Volatility
$h_\star$	14,210	Total firms

Table: Calibrated SIR parameters

Simulated Epidemic Trajectory

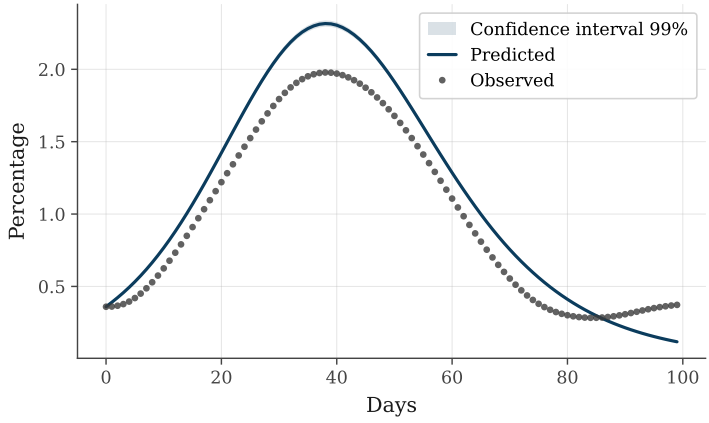

Figure: Total infected subunits (% of total firms). Observed (black) vs simulated mean (red) with 99% CI (blue).

Model captures peak date (day 38) and magnitude (overestimates by 17%).

Infection Probability by Firm Size

k	$\mathbb{P}[\tau_k > 100]$
1	0.794
3	0.500
6	0.249
9	0.129
12	0.060

Table: Probability of no infection in 100 days.

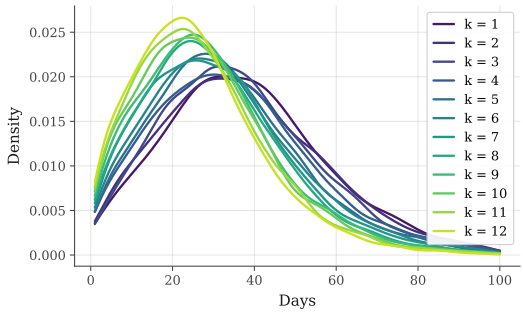

Figure: PDF of infection time by firm size.

Larger firms have higher chance of infection.

Daily Revenue Loss by Firm Size

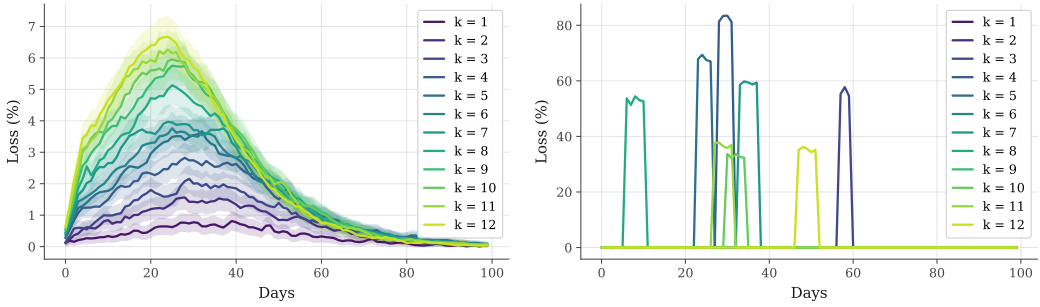

Figure: Average daily loss (left) and one scenario (right) as % of daily revenue.

Larger firms face higher losses due to internal contagion. Losses can spike to 75% in extreme scenarios.

Total Loss Distribution (100 days)

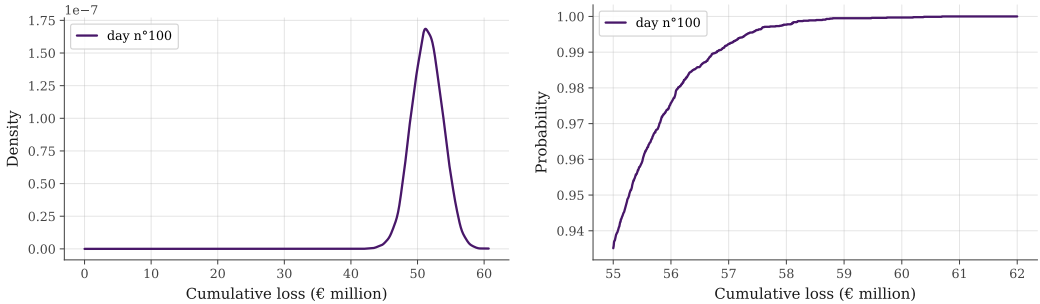

Figure: PDF (left) and right tail (right) of total losses over 100 days.

Total losses can reach €60M ($\approx$ 1.5 days of aggregate revenue). Probability of exceeding €55M is about 6%.

Aggregate Exceedance Probability (AEP)

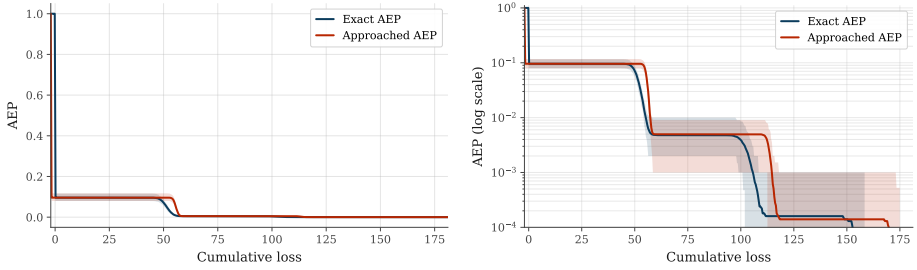

Figure: AEP curves: exact (blue) and approximated (red).

With on average 0.1 episodes per 100 days, the distribution has jumps at multiples of the single-episode loss distribution; extreme losses reach €180M ($\approx$ 4 days of revenue).

Conclusion

Key contributions

- Stochastic multigroup SIR + granular firm growth; captures heavy-tailed sizes, size-dependent contagion, environmental randomness.
- Attack arrivals via Cox (external) + Bernoulli (internal); calibrated on LockBit data.
- Larger firms: higher infection probability and larger losses (internal contagion).
- AEP as an insurer risk measure; extreme losses reach several days of aggregate revenue.

Future work

- SIRS for reinfection; severity dependent on epidemic dynamics or cyber investments.
- Network structures between firms; pricing and risk-transfer; sector-specific calibration.

Thank you for your attention.

Questions?

- **Existence & Uniqueness:** unique global non-negative solution adapted to the filtration generated by the stochastic parameters, $(S_t, I_t, R_t) \in \mathbb{R}_+^{3K},\ \forall t \geq 0$.
- **Basic Reproduction Number:**

$$\mathcal{R}_{k,t} = \frac{\beta_{k,t}}{\gamma_{k,t}} \sum_{i=1}^{K} i \sum_{m=i}^{K} \frac{m S_{m,t}}{N_0} b_{m,i}.$$

- **Peak:** $\mathfrak{I}_{\max} = \max_{t \geq 0} \sum_{k=1}^{K} k I_{k,t}$.
- **Prevalence:** $\mathfrak{P}_t = \frac{1}{N_0} \sum_{k=1}^{K} k R_{k,t}$.

If $\max_k \mathcal{R}_{k,t} < 1$ then $\frac{d}{dt} \log \mathfrak{I}_t < 0$: total infected decreases (locally exponentially).

- $N^0_{ij,t}$ models infections from outside the firm (other firms / external attackers).
- Once a subunit of firm i is externally infected at τ_i, still-susceptible subunits may be infected internally with probability a_{τ_i} (in-firm rate from the SIR model).
- Internal infection is the Bernoulli variable U_{ij} ($U_{ij} = 1 \rightarrow$ infected at τ_i).
- Captures the empirical fact that infections spread within an organization after an initial breach.

Backup — Peak Distribution by Firm Size

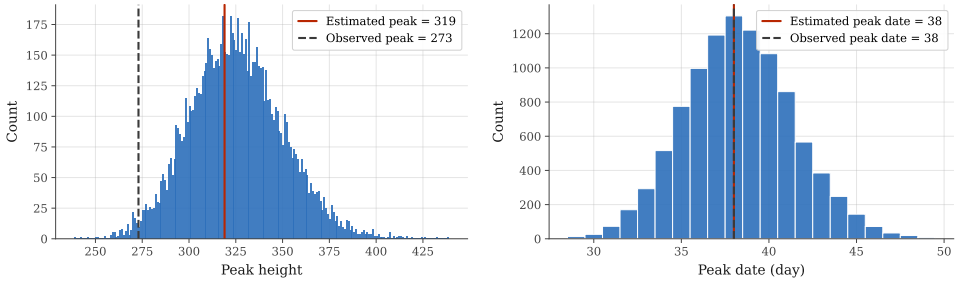

Figure: Histograms of peak height (left) and peak date (right).

Backup — Daily Loss Distribution

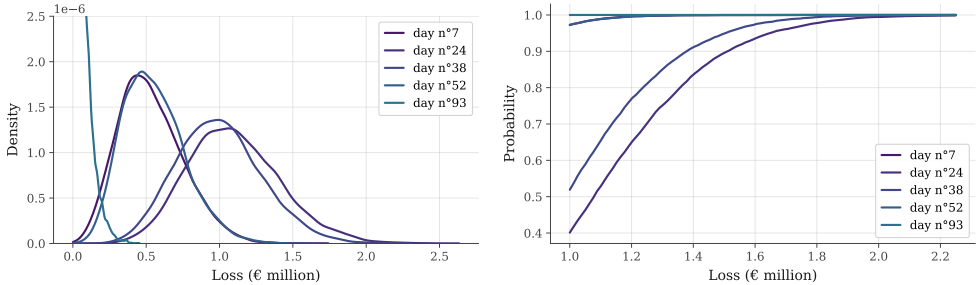

Figure: PDF (left) and right tail of CDF (right) of daily losses (in €M) for sub-portfolio of 621 firms ($K_i \geq 2$).

Around the epidemic peak (day 24), the distribution widens and tails become heavier.

References I

Acemoglu, D., Malekian, A., and Ozdaglar, A. (2016).
Network security and contagion.
Journal of Economic Theory, 166:536–585.

AXA Group (2025).
Future risks report 2025.
Accessed: 2026-02-10.

Axtell, R. L. (2001).
Zipf distribution of us firm sizes.
science, 293(5536):1818–1820.

Baksy, A., Caratelli, D., and Olson, L. M. (2025).
Cyberattacks and firm size: The vulnerability of mid-size firms.
Office of Financial Research, The OFR Blog.

Bessy-Roland, Y., Boumezoued, A., and Hillairet, C. (2021).
Multivariate Hawkes process for cyber insurance.
Annals of Actuarial Science, 15(1):14–39.

Boumezoued, A., Cherkaoui, Y., and Hillairet, C. (2025).
Cyber risk frequency modelling using Hawkes processes: Calibration on attack and vulnerability data.
https://hal.science/hal-05305048v1.

Edwards, B., Hofmeyr, S., and Forrest, S. (2016).
Hype and heavy tails: A closer look at data breaches.
Journal of Cybersecurity, 2(1):3–14.

References II

Eling, M. and Loperfido, N. (2017).
Data breaches: Goodness of fit, pricing, and risk measurement.
Insurance: Mathematics and Economics, 75:126–136.

European Central Bank (2025).
Cyber resilience stress testing: A macroprudential approach.
Macroprudential Bulletin, Issue 22.

Fahrenwaldt, M. A., Weber, S., and Weske, K. (2018).
Pricing of cyber insurance contracts in a network model.
ASTIN Bulletin: The Journal of the IAA, 48(3):1175–1218.

Farkas, S., Lopez, O., and Thomas, M. (2021).
Cyber claim analysis using generalized Pareto regression trees with applications to insurance.
Insurance: Mathematics and Economics, 98:92–105.

Gabaix, X. (1999).
Zipf's law and the growth of cities.
American Economic Review, 89(2):129–132.

GuidePoint Security (2025).
GRIT 2025 Ransomware & Cyber threat report.
Technical report, GuidePoint Security.
Accessed: 2025-08-15.

Hillairet, C. and Lopez, O. (2021).
Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models.
Scandinavian Actuarial Journal, 2021(8):671–694.

References III

Hillairet, C., Lopez, O., d'Oultremont, L., and Spoorenberg, B. (2022).
Cyber contagion: impact of the network structure on the losses of an insurance portfolio.
Insurance: Mathematics and Economics.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., and Stulz, R. M. (2021).
Risk management, firm reputation, and the impact of successful cyberattacks on target firms.
Journal of Financial Economics, 139(3):719–749.

Langlois-Berthelot, J., Gaie, C., and Lebraty, J.-F. (2024).
Epidemiology inspired cybersecurity threats forecasting models applied to e-government.
In *Transforming Public Services—Combining Data and Algorithms to Fulfil Citizen's Expectations*, pages 151–174. Springer.

Moran, J., Secchi, A., and Bouchaud, J.-P. (2024).
Revisiting granular models of firm growth.
Working paper.

Murray, W. (1988).
The application of epidemiology to computer viruses.
Computers & Security, 7(2):139–145.

Stanley, M. H., Amaral, L. A., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A., and Stanley, H. E. (1996).
Scaling behaviour in the growth of companies.
Nature, 379(6568):804–806.

Wyart, M. and Bouchaud, J.-P. (2003).
Statistical models for company growth.
Physica A: Statistical Mechanics and its Applications, 326(1-2):241–255.