# Self-Organization of R&D Search in Complex Technology Spaces

Gerald Silverberg*

based on collaboration with Bart Verspagen**

* IIASA and MERIT, Maastricht University

** ECIS, Eindhoven University of Technology

# Stylized Facts About Technological Change

1. Technical change is cumulative: new technologies build on each other.

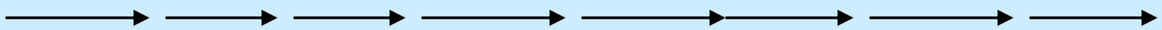
2. Technical change follows relatively ordered pathways, as can be measured ex post in technology characteristics space (see the work of Sahal, Saviotti, Foray and Gruebler, etc.). This has led to the positing of *natural trajectories* (Nelson), *technological paradigms* (Dosi), and *technological guideposts* (Sahal).

3. The arrival of innovations appears to be stochastic, but more highly clustered than Poisson (overdispersion).

4. The ‘size’ of an innovation is drawn from a highly skewed and possibly fat-tailed distribution.

# Standard Perspective on Technical Change: One Dimensional Search

- R&D as stepping up a quality ladder:



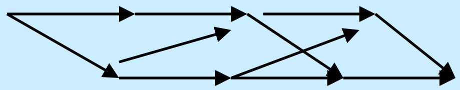
- Each rung represents an improvement in performance by a factor $\gamma$ attainable with certainty
- Realizing each step is modelled as a Poisson-process patent race with exponentially distributed waiting times:

  Prob(innovation made in elapsed time t) = 1 – exp(-Rt),

  where R is a measure of R&D investment intensity per period.
- Patent races are mutually exclusive, winner-take-all contests with duplicative R&D effort in parallel, no synergies or spillovers

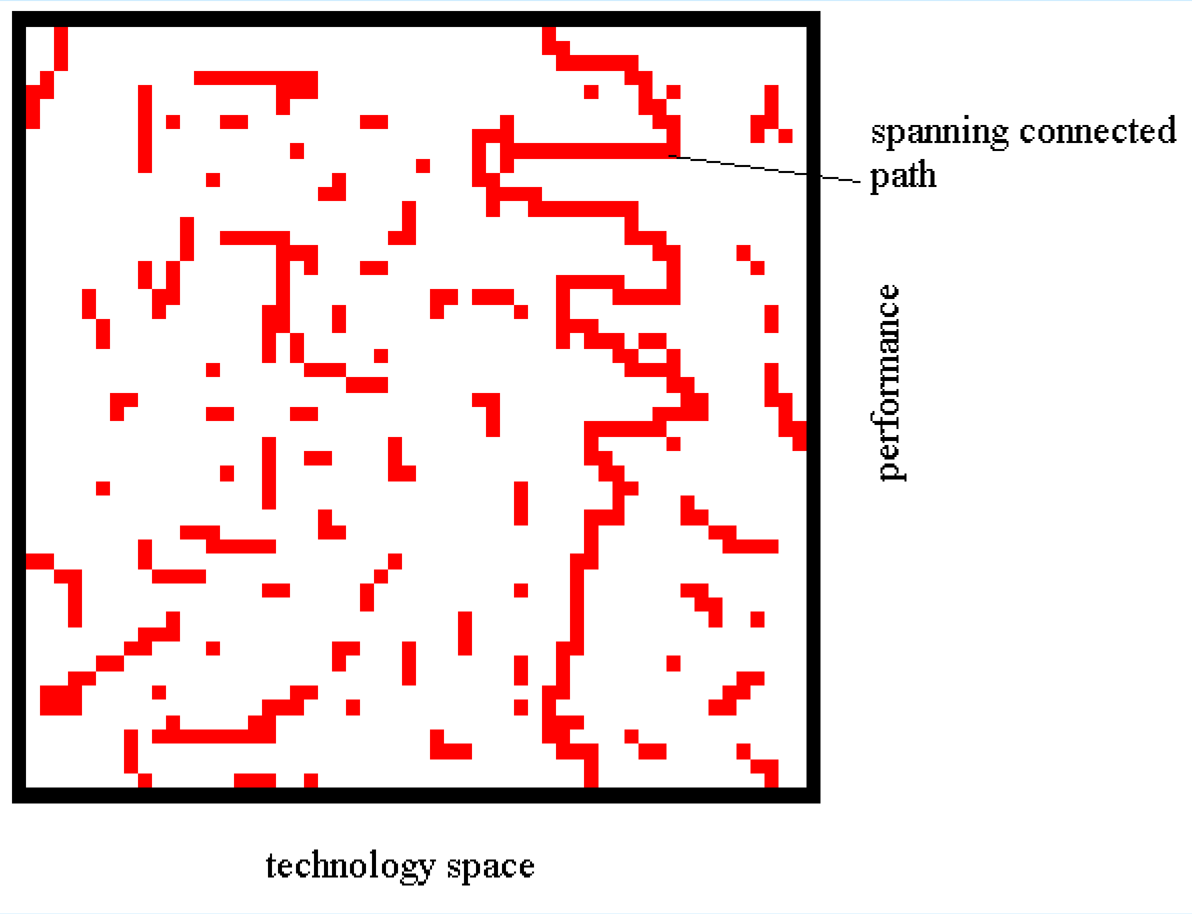
# Complex Technology Spaces: Rugged and Dimension ≥ 2

- Consider technologies as being nodes in a graph, with (possibly directed) edges representing minimal innovative steps in the search process

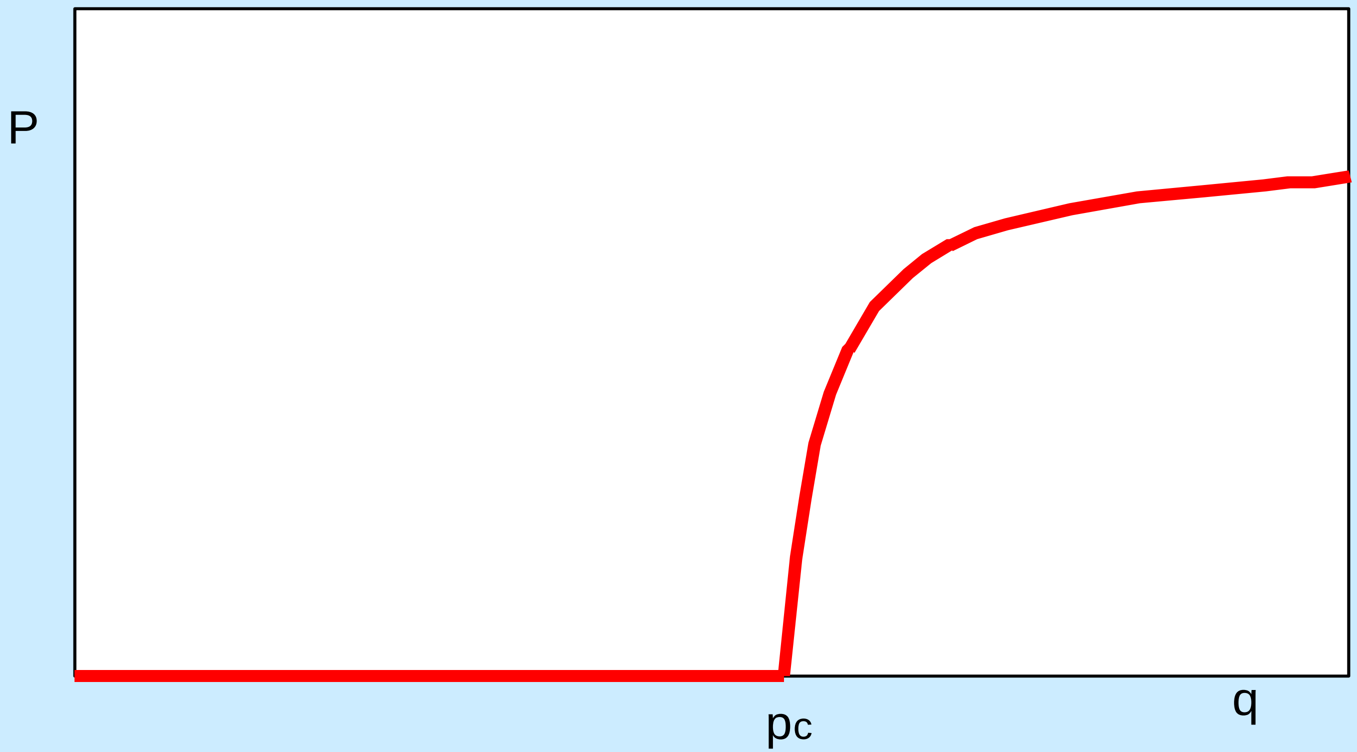


- The nodes can differ in difficulty of realization (extreme case: possible/impossible) and in payoffs
- If the difficulty of realization is a random variable (extreme case: iid between nodes) the space becomes rugged
- Simplest generalization of linear quality ladder: two dimensional lattice
- Ruggedness would seem to stymie innovation by making unbounded paths through the space practically impossible
- However, the addition of a second (or higher) dimension to the space means that percolation will occur for certain parameter values, making a (highly irregular) infinite spanning cluster a possible technological trajectory through the space

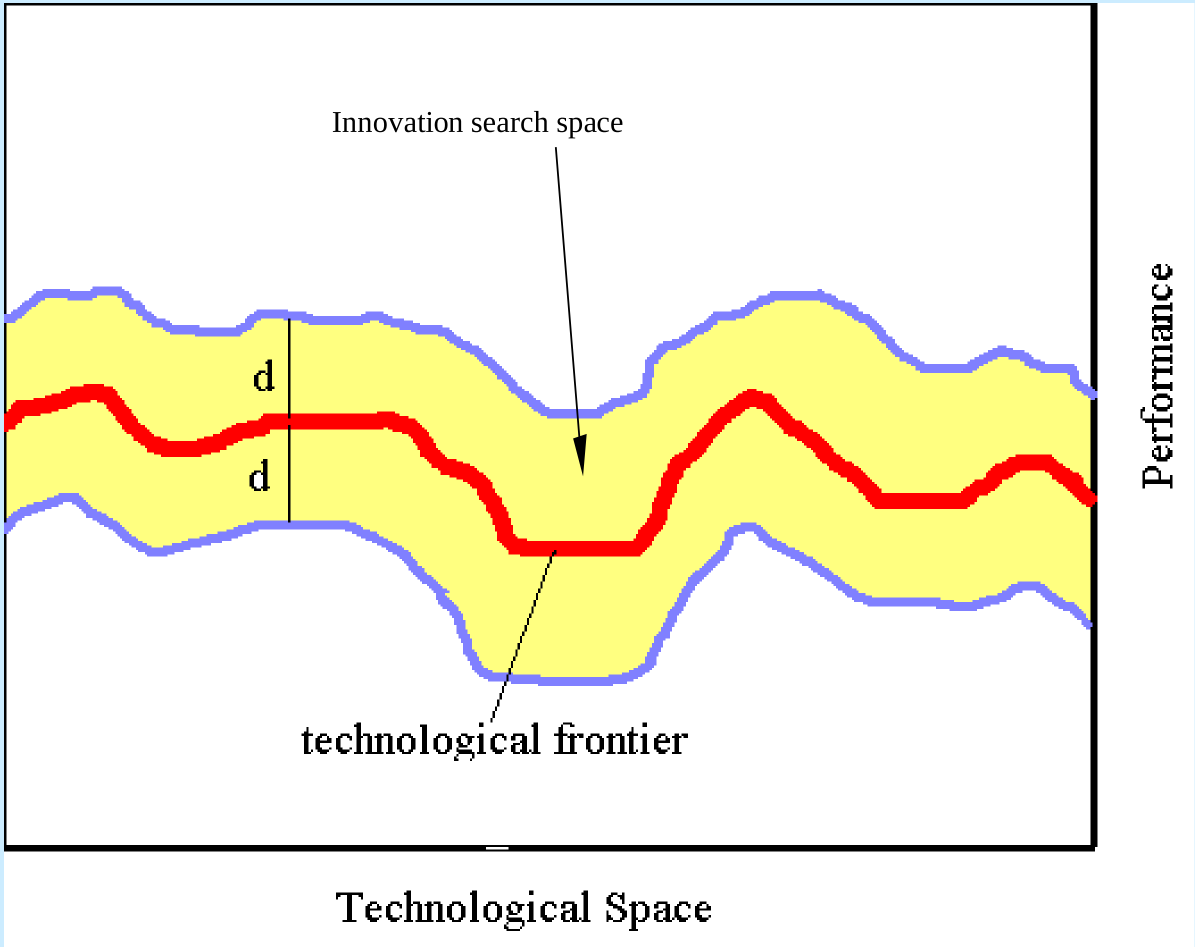
**Percolation diagram in technology-performance space. Lattice sites are filled at random. A site is viable when it connects to the baseline.**



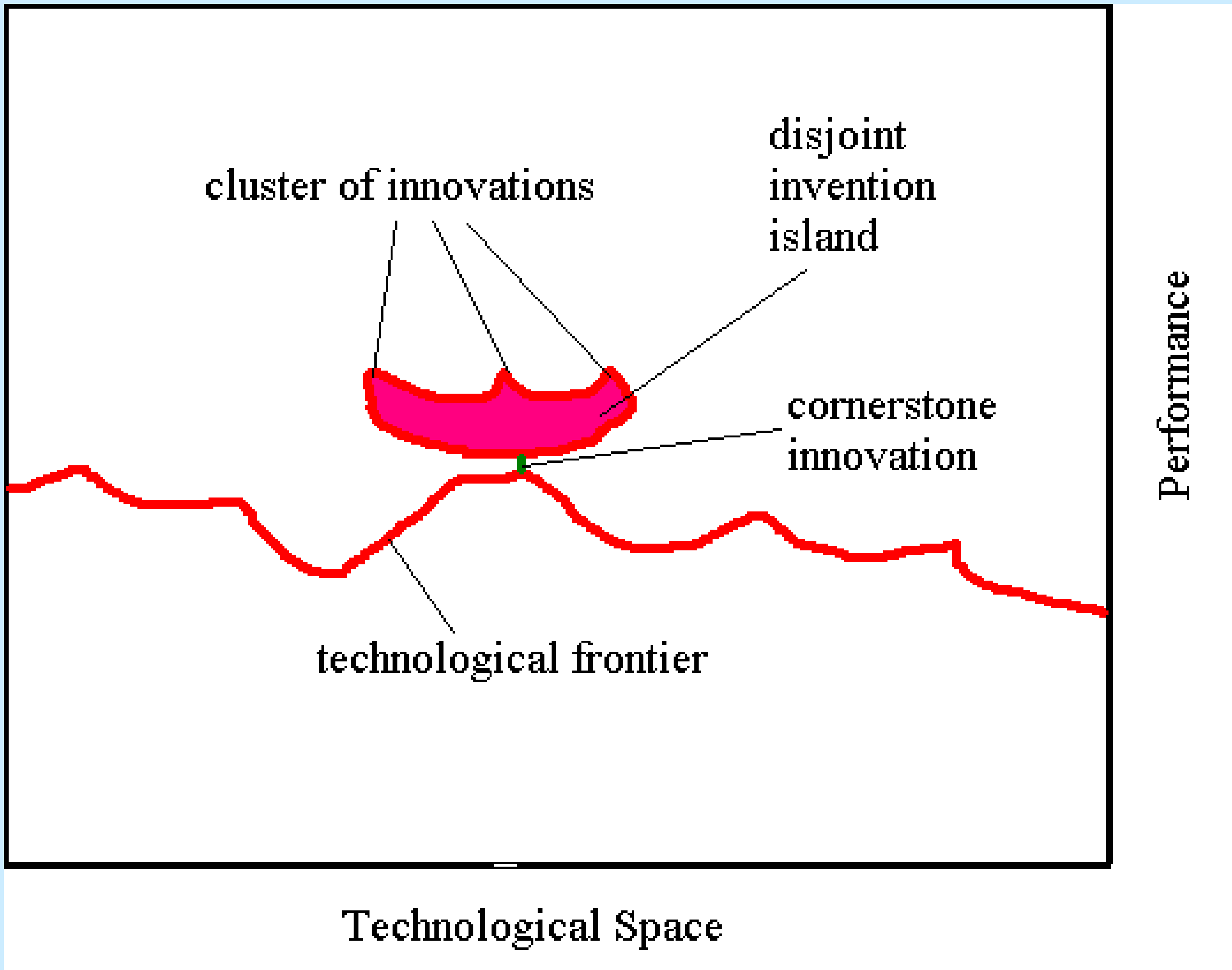
**Probability of a random site being on the infinite cluster *P* as a function of the percolation probability *q***



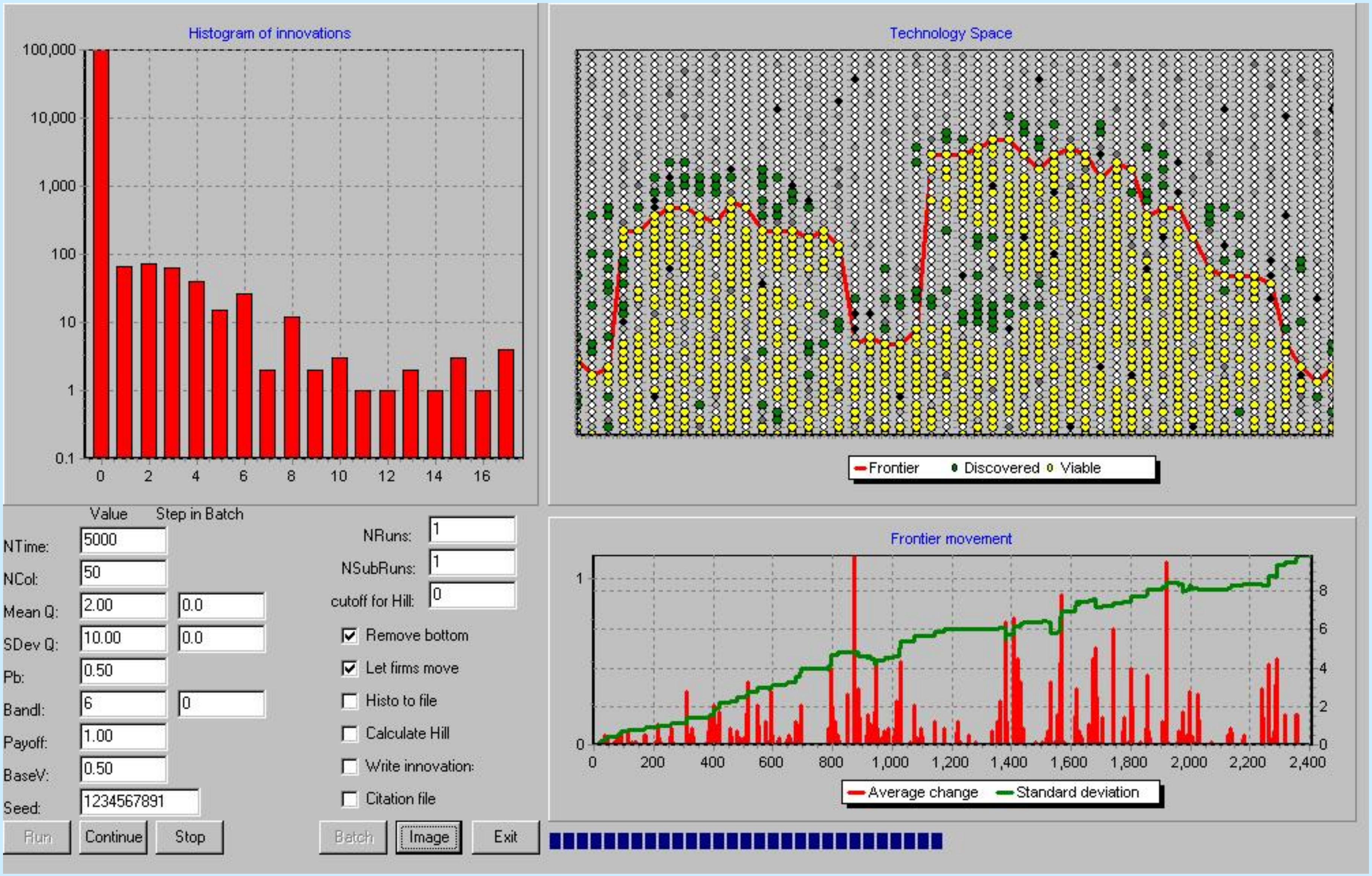
**New innovations are generated with probability p in a region d units above and below the technological frontier.**



**A cluster of simultaneous invention occurs when a disjoint island of invention is suddenly joined to the frontier by a single 'cornerstone' innovation.**



# Screen Shot of Run with Search Radius 6



# What's New in the New (Self-Organizational) Model

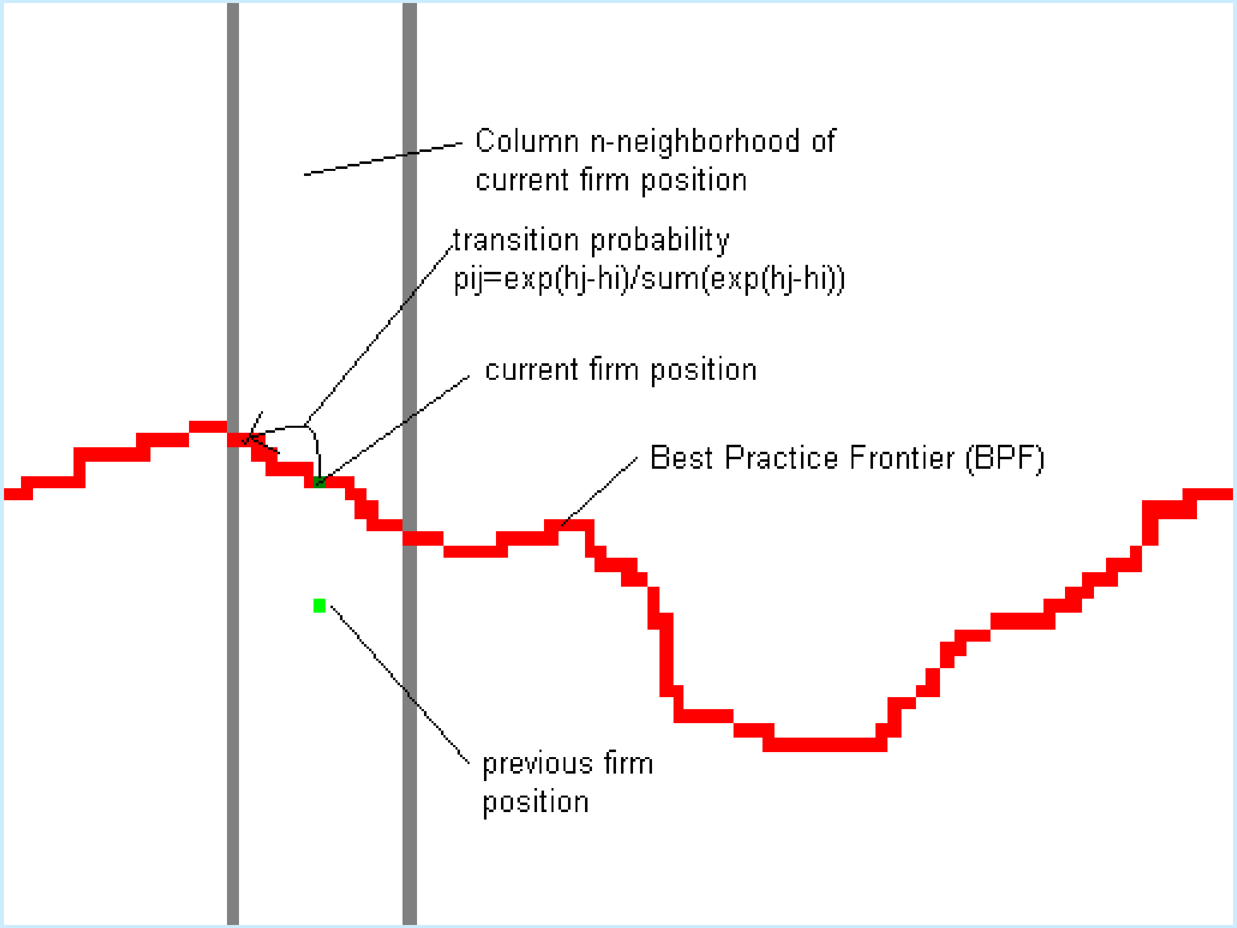
- Toyota principle: "Nothing is impossible".
  Instead of {0, 1} random values for site feasibility, we now draw a 'difficulty' value q from a LN(<q>, σ) distribution. High skewness of LN, however, makes some sites very difficult.

- Firms now 'dig' out the site, with complete technological spillovers between periods: $q_i(t+1) = \max(0, q_i(t) - B\omega)$, where ω is drawn from a uniform distribution on [0, 1).

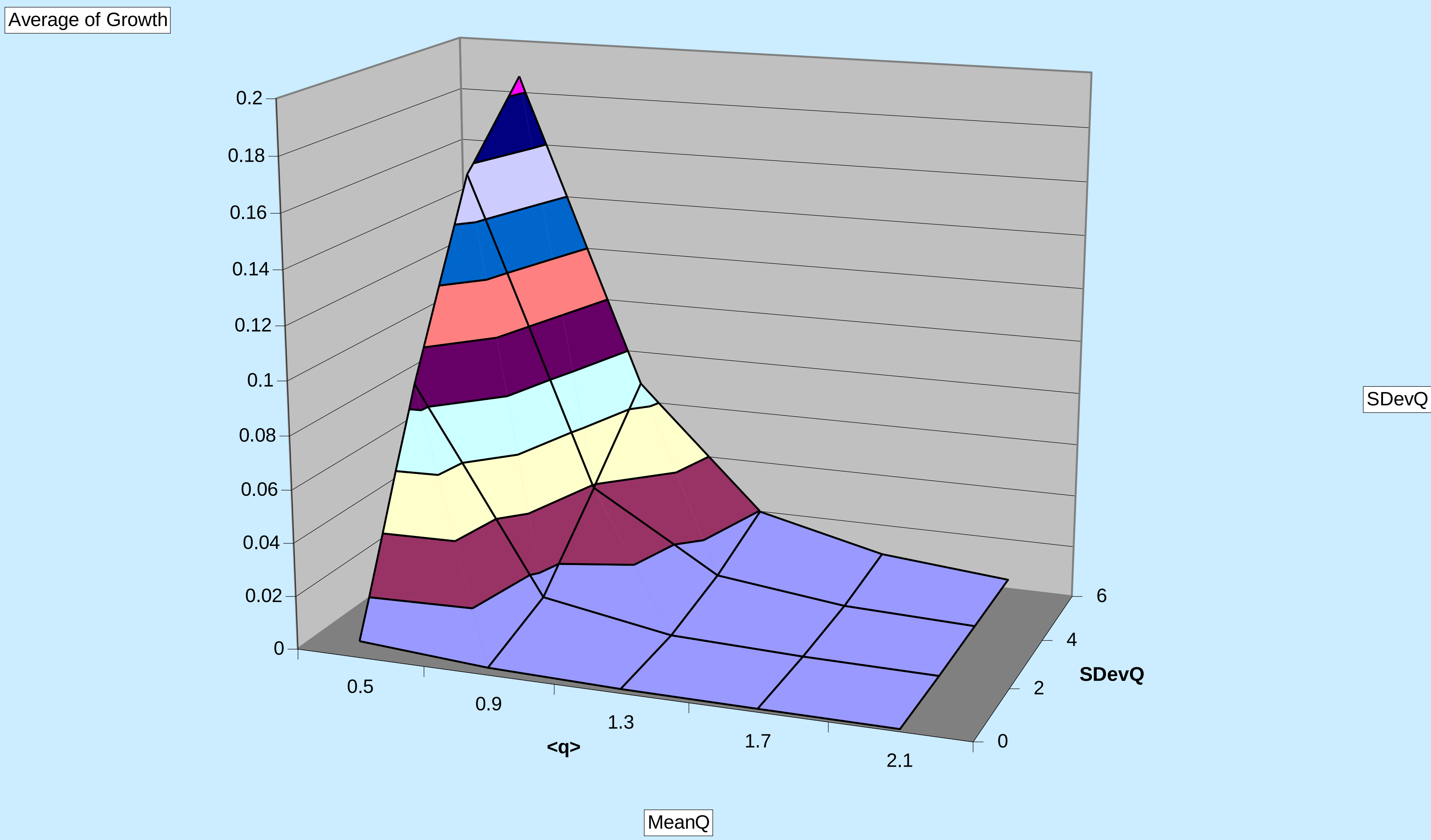
- We now allow myopic 'strategic' firm behavior: firms position themselves within their technological search neighborhoods, moving with highest probability to the site which is highest on the BPF based on multinomial transition probabilities.

- Firms now profit from successful innovations, with their R&D budgets increasing in proportion to their innovation 'harvest'.

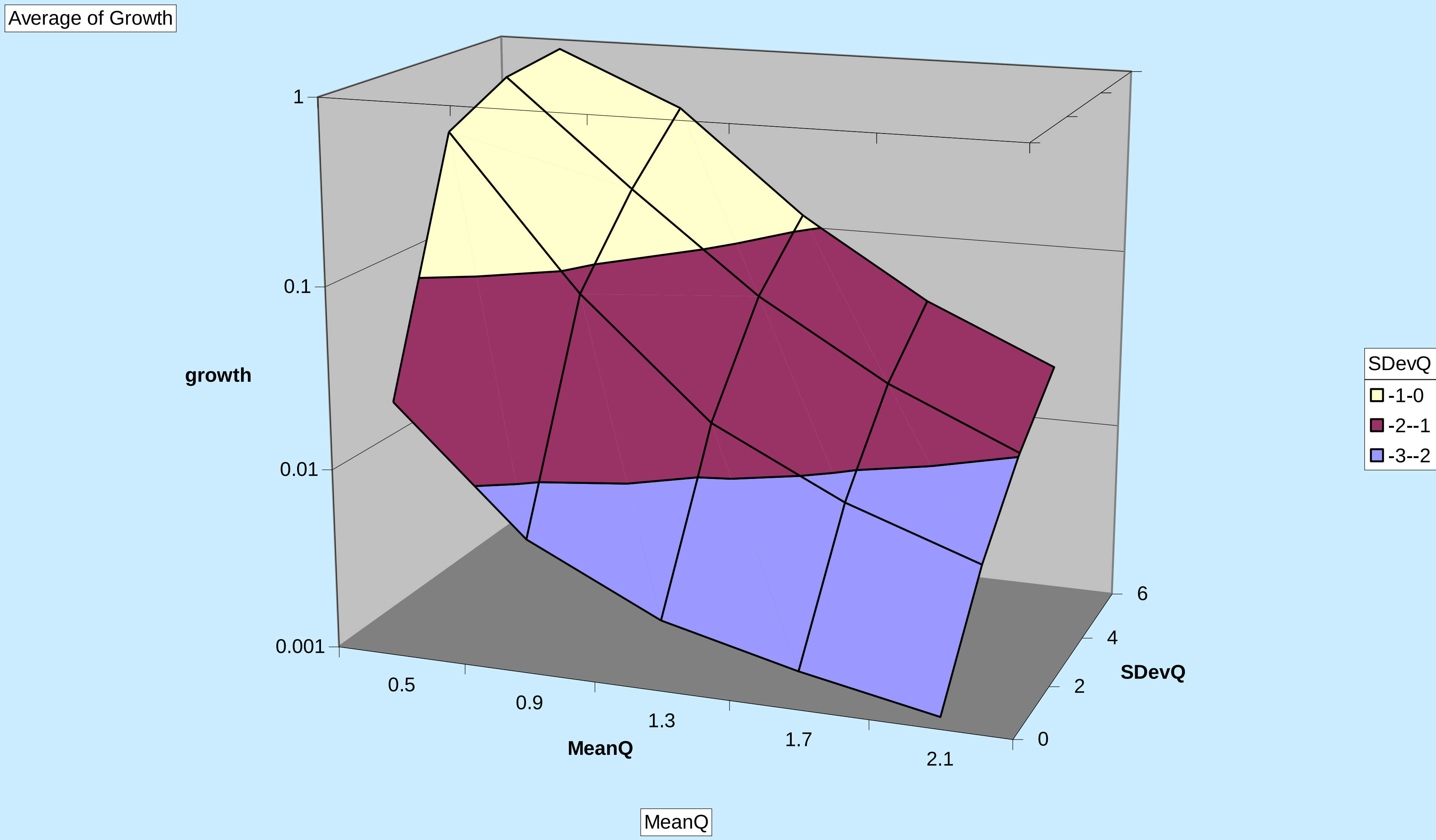
Firms Move by Myopically Searching for Higher Positions within their Column n-Neighborhood on the BPF before Untertaking Next R&D Round



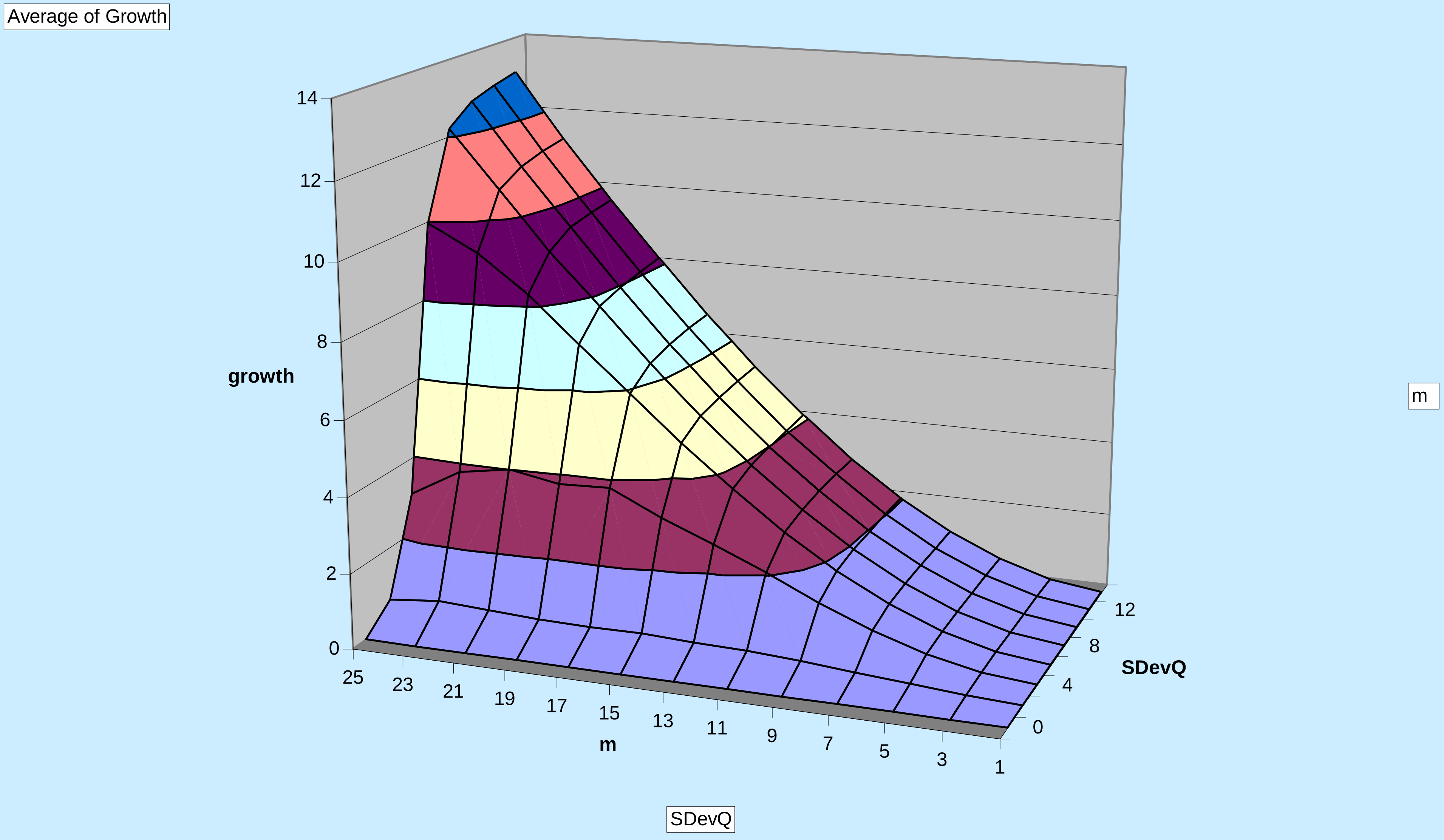
# Average innovation rate as a function of <q> and σ for fixed firms



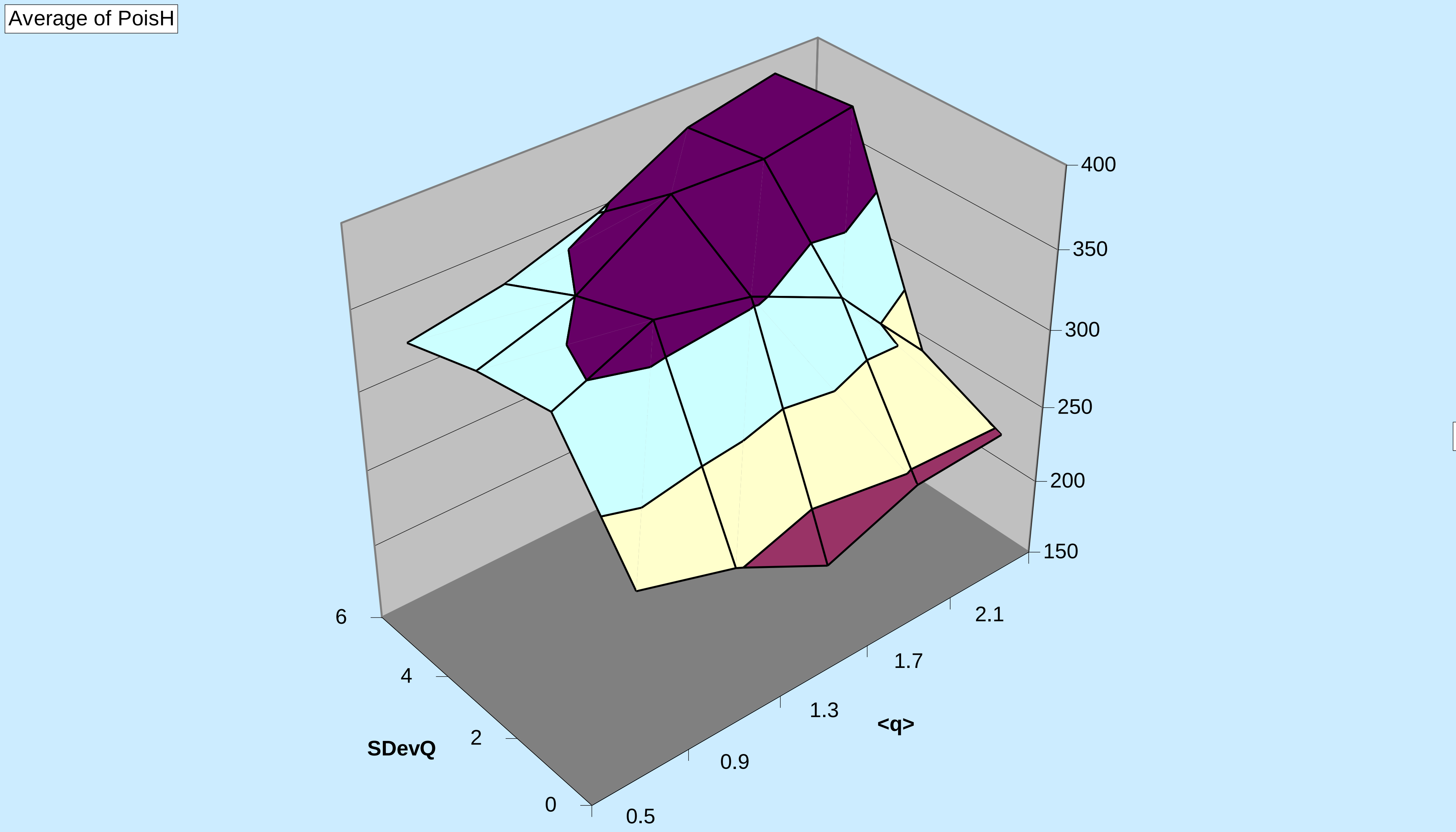
# Average innovation rate as a function of <q> and σ, moving firms



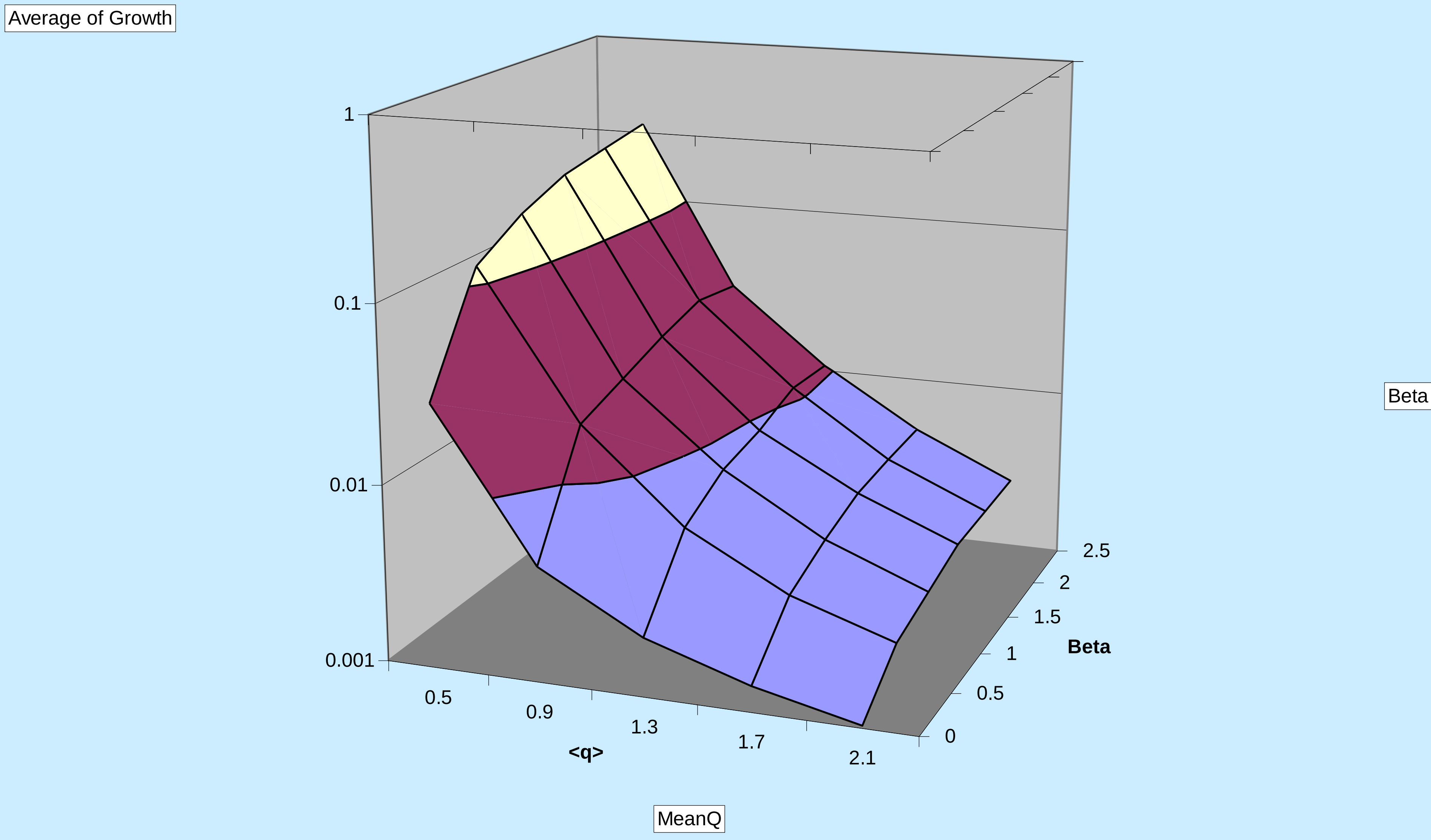
# Average innovation growth rate as a function of search radius m and σ, fixed firms



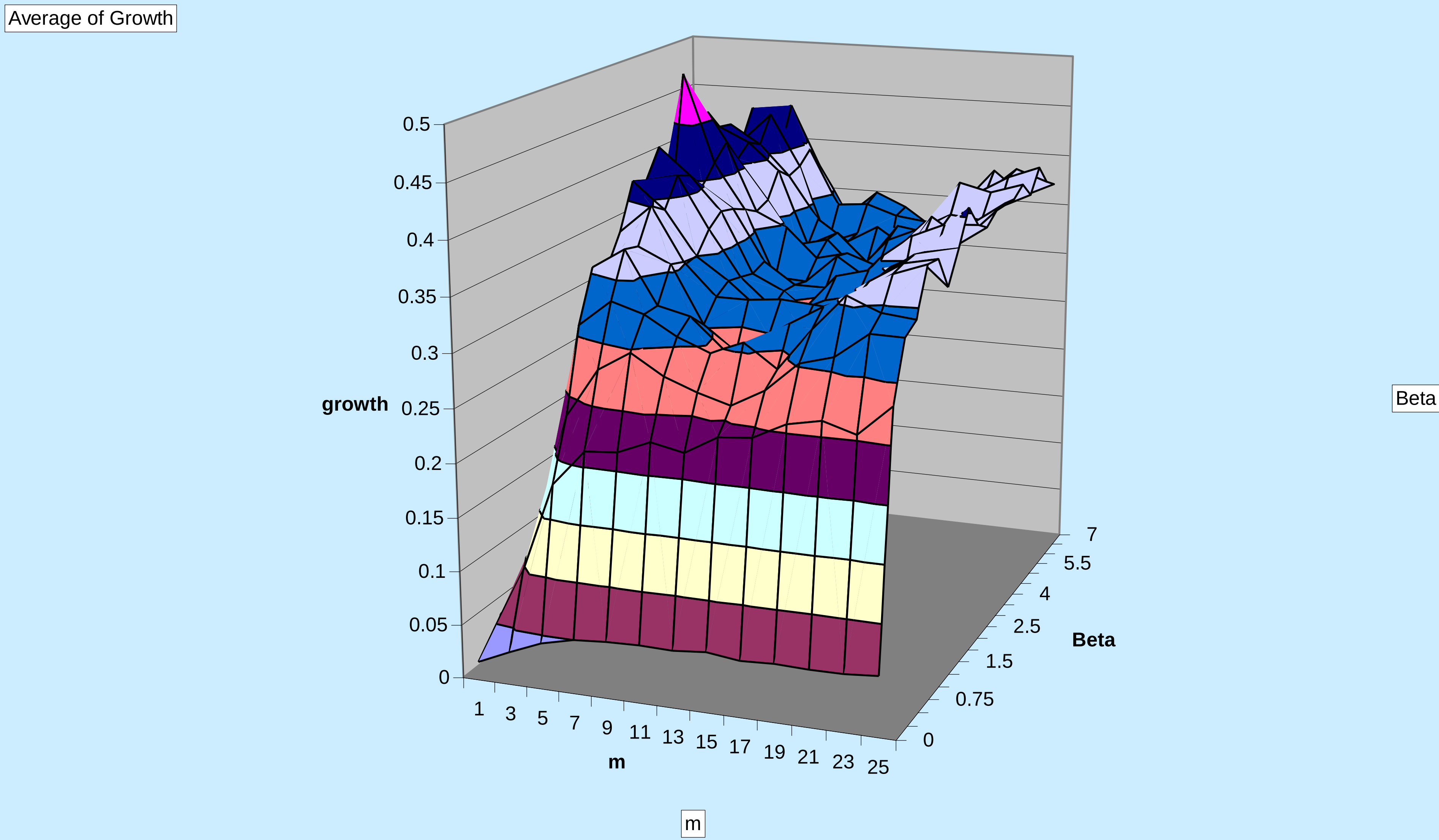
# Clustering as a function of <q> and σ



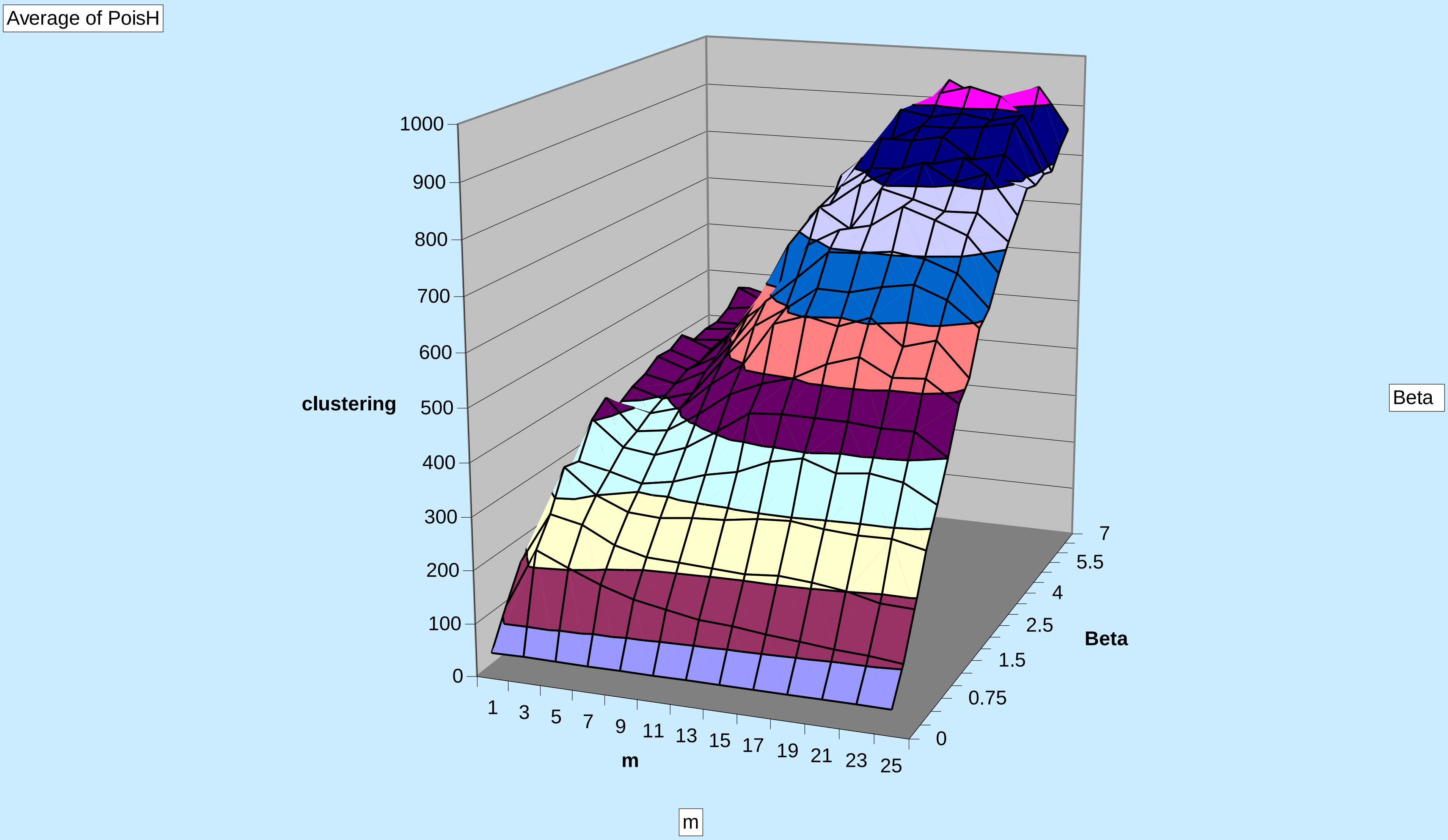
# Average innovation rate as function of <q> and ‘rationality’ β, moving firms



# Average innovation rate as a function of search radius m and 'rationality' β, moving firms



# Clustering index as a function of search radius m and ‘rationality’ β, moving firms



# Time Series of Clustering Index

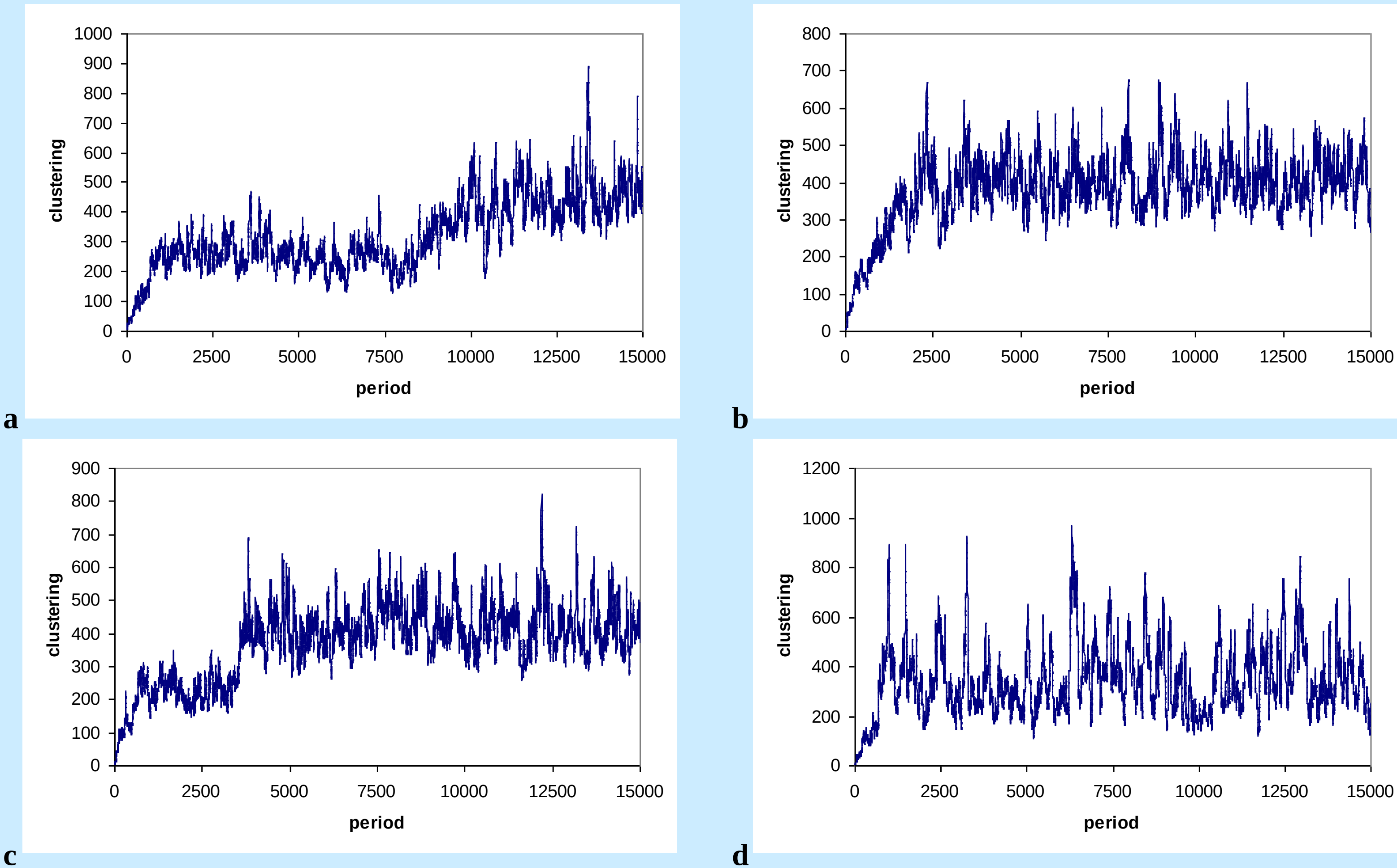
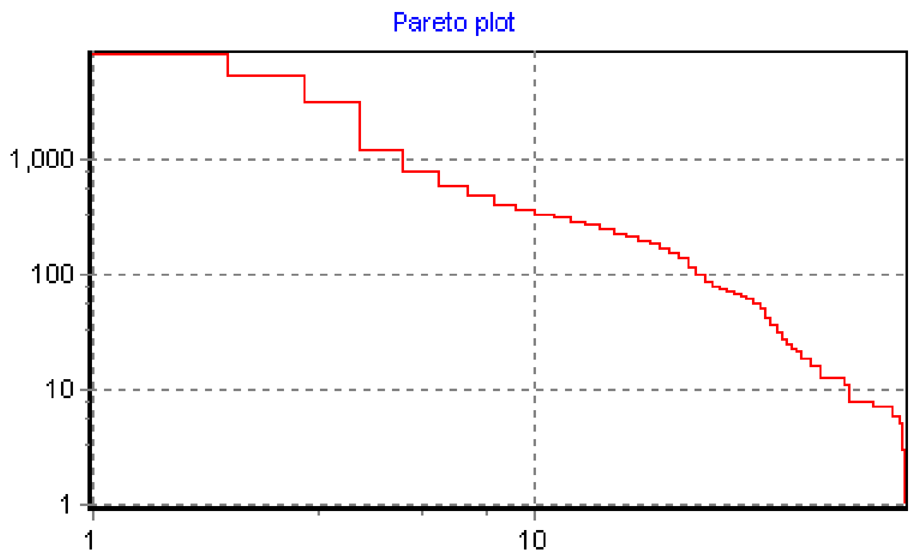
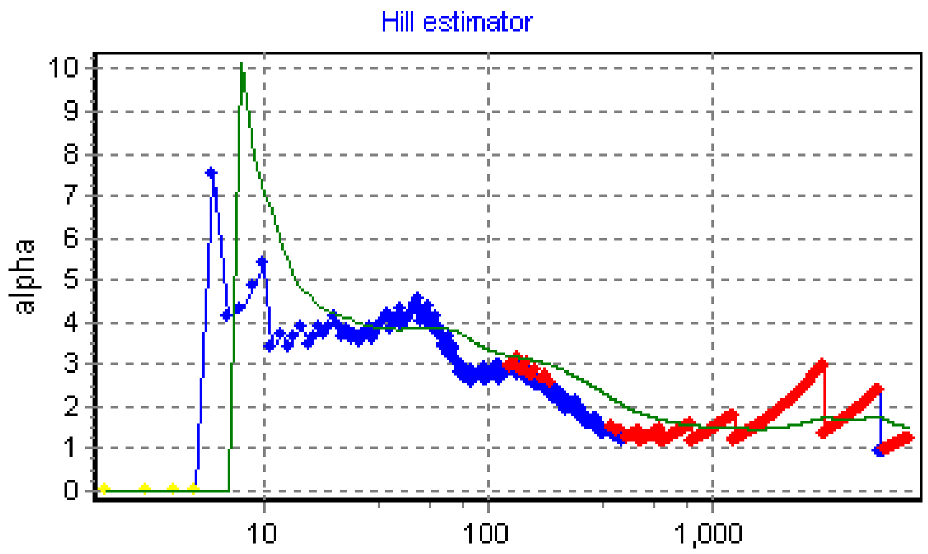


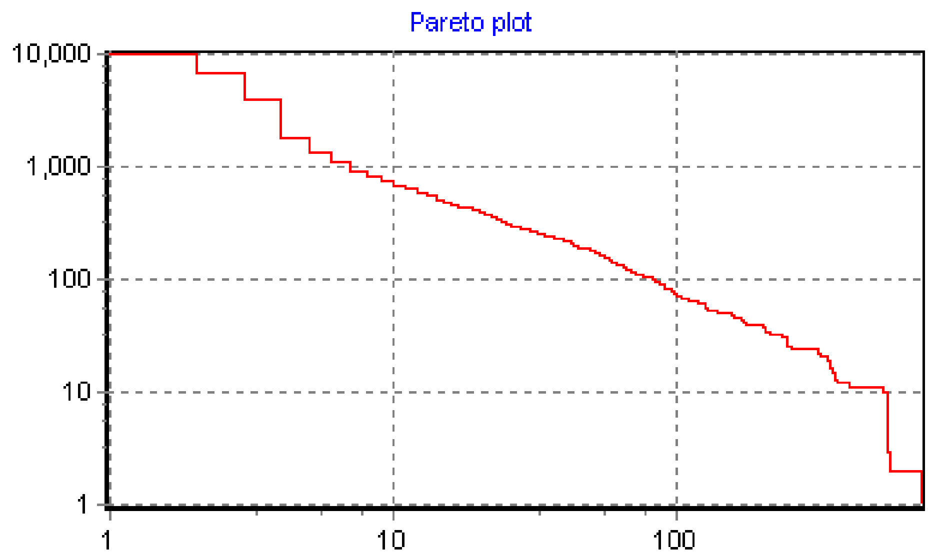

**Figure 1 Clustering index for four runs with $<q>=0.5$, $\pi=1$ and: (a) $\sigma=1$, $m=3$; (b) $\sigma=2$, $m=3$; (c) $\sigma=4$, $m=3$; and (d) $\sigma=1$, $m=10$. Bimodality appears to be present in (a) and (c).**

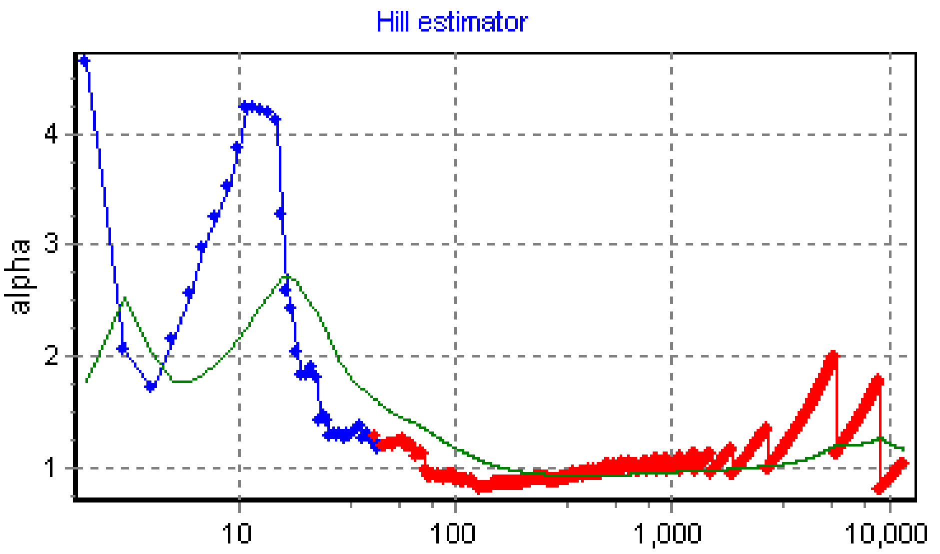
# Innovation Size Distribution, Fixed Firms, <q>=0.5, σ=1, m=3, π=1



# Innovation size distribution, moving firms, <q>=0.5, σ=1, m=3, π=1



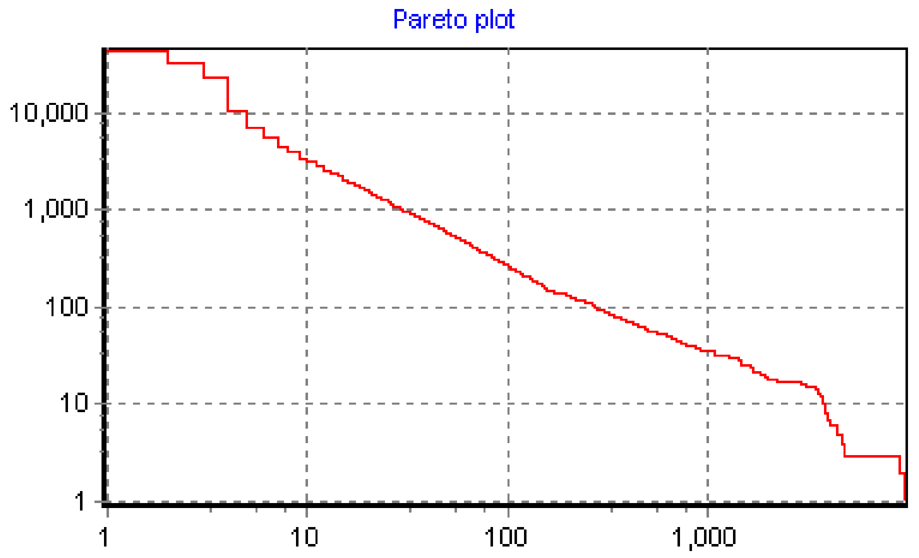



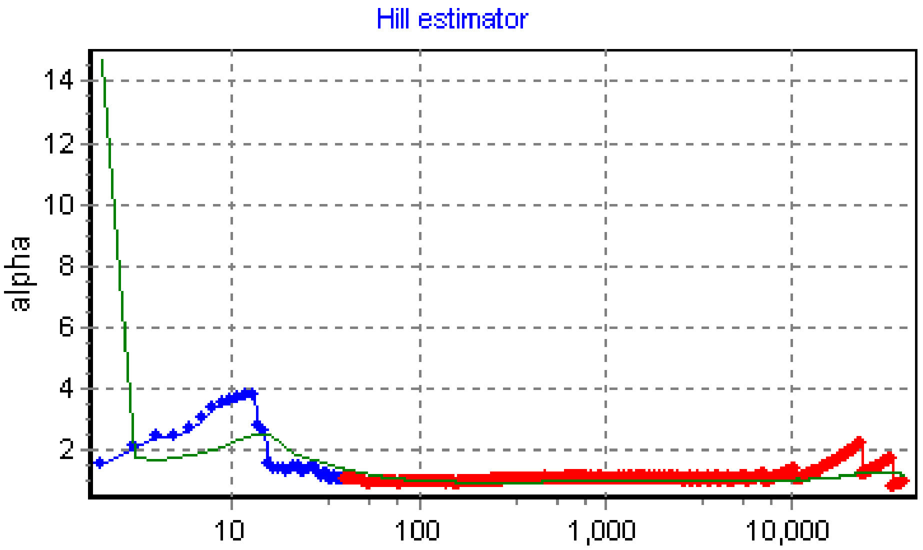
➔Amazing Fact!➔

These innovation size distributions for moving firms look the same ($\alpha \approx 1$, scaling over more than two decades) for a wide range of parameter values.

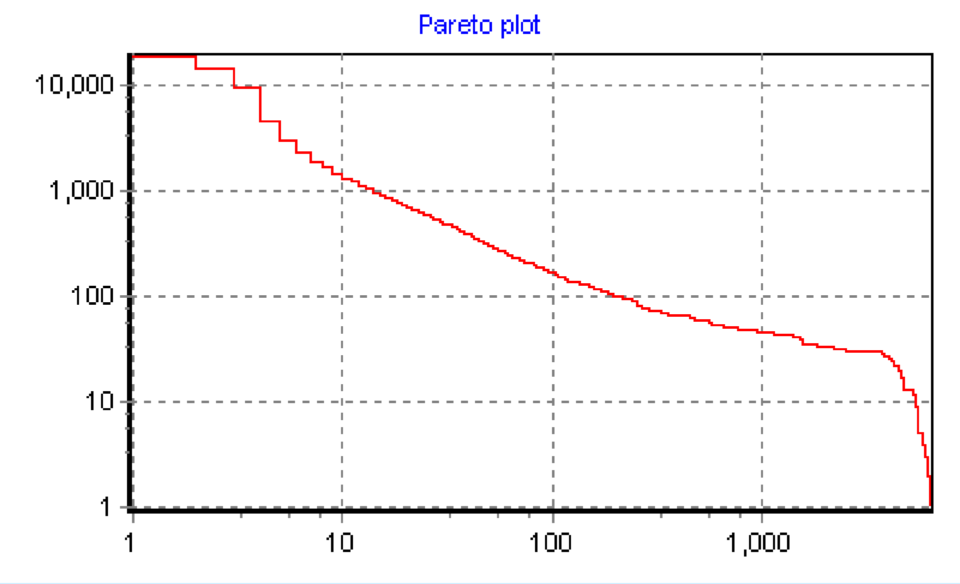
SOC??

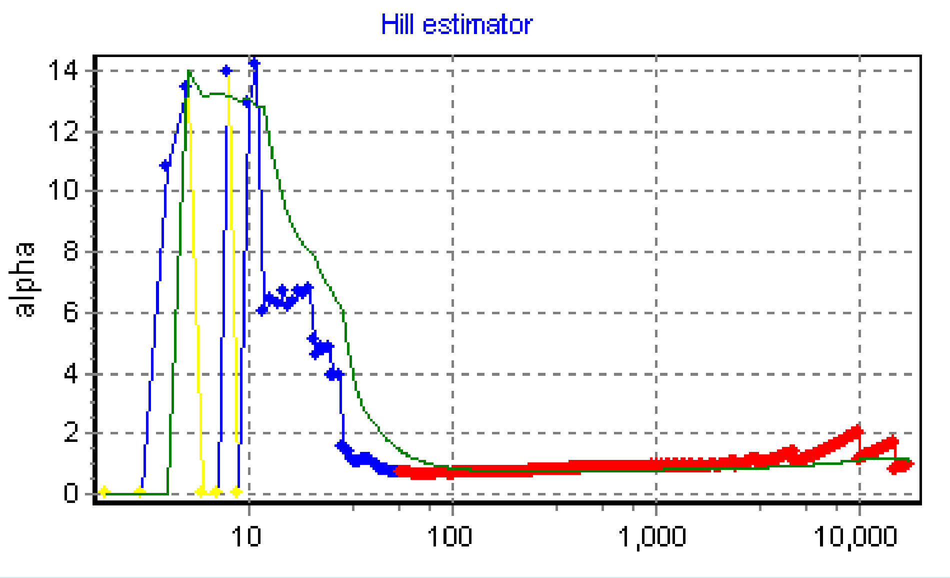
# Innovation size distribution for same parameters except σ=2



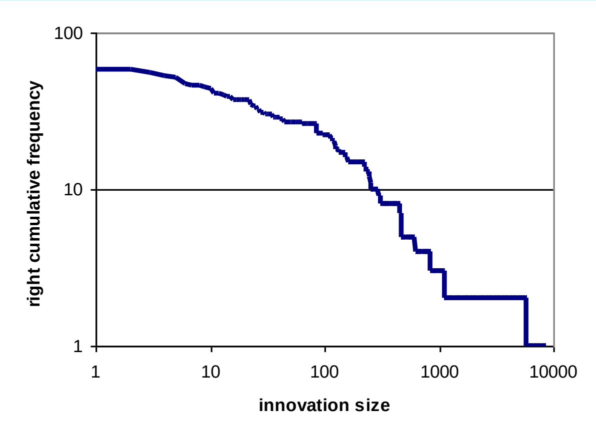



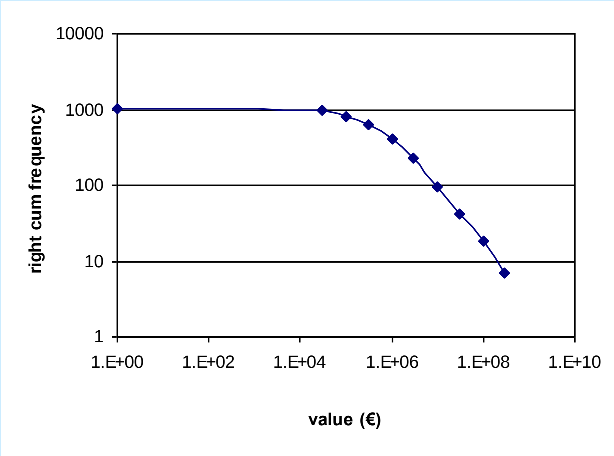
# Innovation size distributions for same parameters except β=2

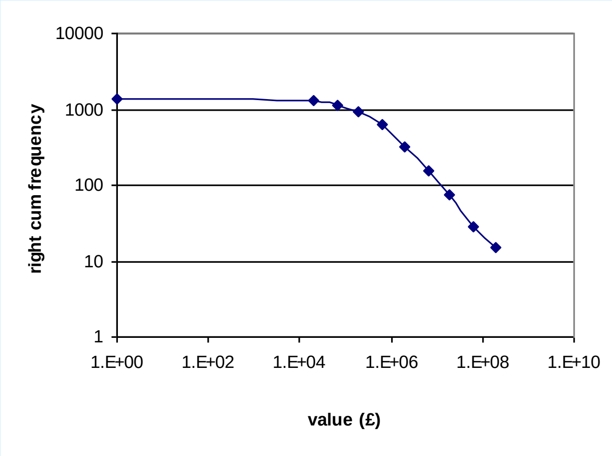


# Innovation Size Distributions (Pareto Plots) Based on Monetary Value




Harvard patent license fees (left), NL patent valuation survey (middle), UK patent valuation survey (right)

# The Hill Estimator

Placing the $n$ observations $X_i$ in descending order and denoting the resulting rank-order statistics by $X_{[i]}$, $X_{[1]} \geq X_{[2]} \geq \ldots \geq X_{[n]}$, the Hill estimator is defined as:

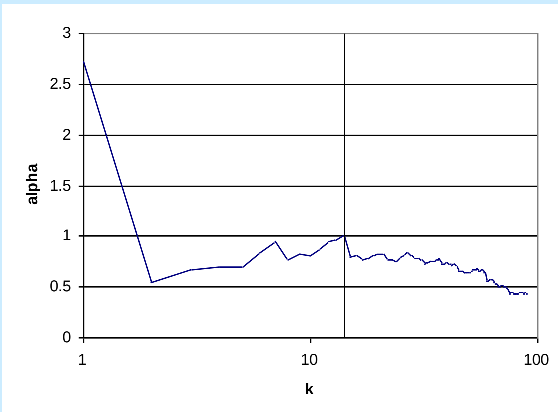
$$H(k,n) = 1/k \sum^{k} (\ln X_{[i]} - \ln X_{[k+1]}).$$

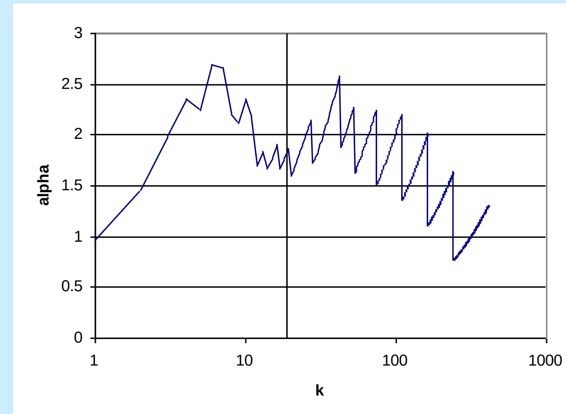
## ML Estimator on Grouped Data

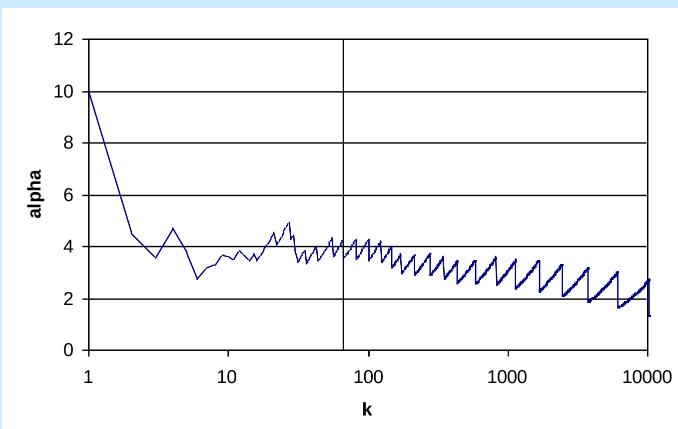
Data are counts of number of observations *ni* in bins [Lc, L1), [L1, L2), … , [Lm-1, Lm), [Lm, ∞). The log likelihood function is then:

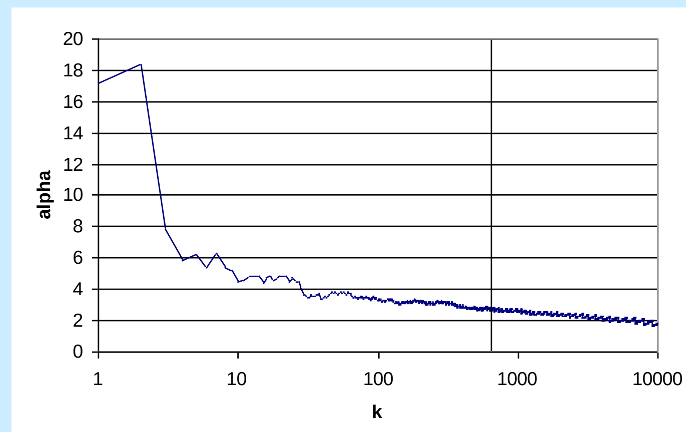
$$\log L(\alpha, L_c) = \sum_{i=c}^{m-1} n_i \log[\,(L_i / L_c)^{-\alpha} - (L_{i+1} / L_c)^{-\alpha}\,] + n_m \log(L_m / L_c)^{-\alpha},$$

## The Dress/Kaufman Tail Determination Using the Hill Diagram



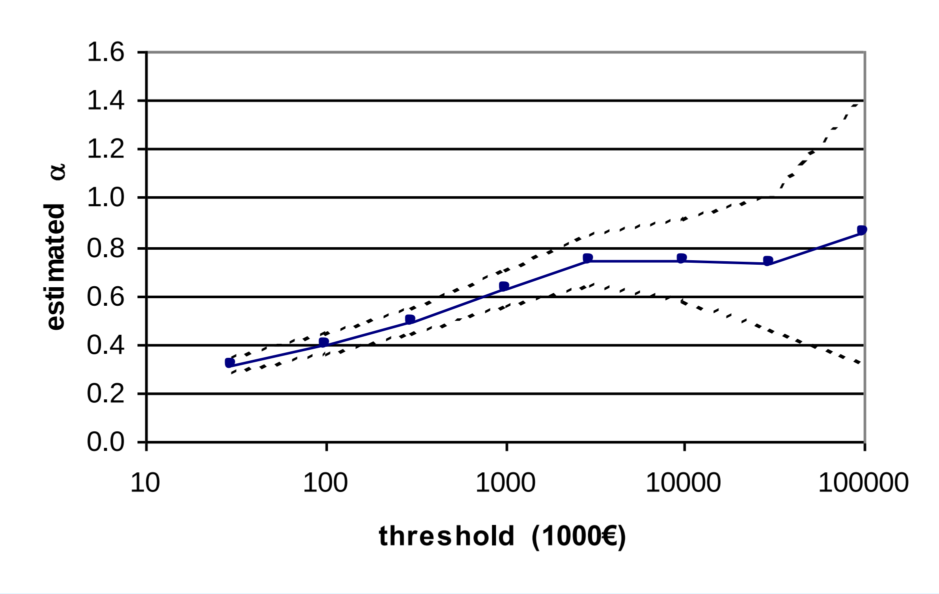



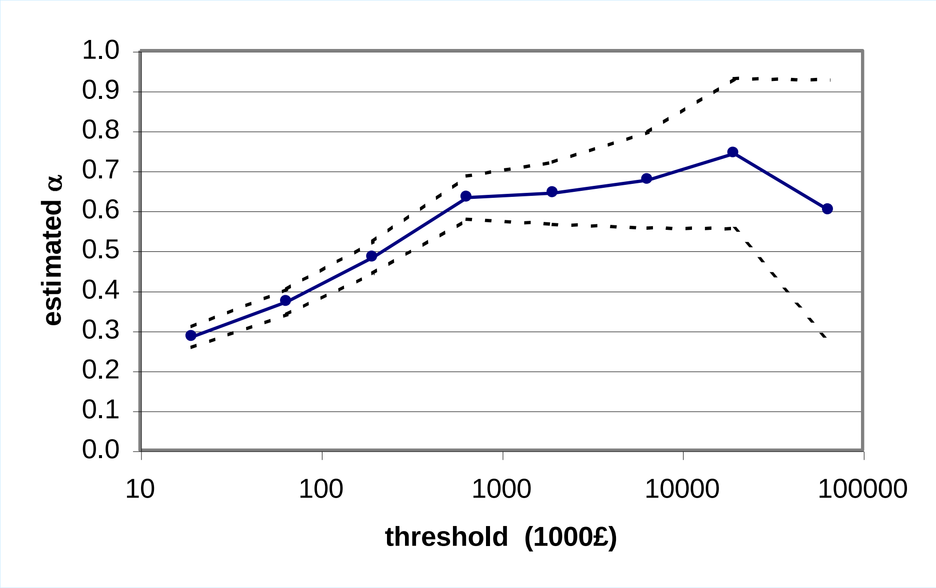




Hill plots and tail cutoff values: Harvard (upper left), Trajtenberg (upper right), EPO (lower right), US (lower left).

# Determining the Tail of the PatVal Datasets



# The Tail, its Fatness, and its Width

**(estimated using a data-driven Hill estimator, Silverberg and Verspagen 2004)**

**Table I. Estimates of tail index α by the Drees/Kaufmann (DK) and Danielsson/deVries (DV) methods**

| Dataset | α | confidence interval | | threshold | k* | n |
|---|---|---|---|---|---|---|
| Harvard DK | 1.010 | 0.480 | 1.537 | $230K | 14 | 100 |
| Harvard DV | 0.823 | 0.462 | 1.183 | $123K | 20 | 100 |
| Trajtenberg DK | 1.864 | 1.026 | 2.701 | 9 | 19 | 456 |
| Trajtenberg DV | 2.333 | 1.591 | 3.074 | 7 | 38 | 456 |
| EPO1989 DK | 3.542 | 2.694 | 4.390 | 20 | 67 | 33499 |
| EPO 1989 DV | 3.371 | 3.207 | 3.535 | 7 | 1615 | 33499 |
| US 1989 DK | 2.718 | 2.509 | 2.927 | 36 | 650 | 50687 |
| US 1989 DV | 2.689 | 2.457 | 2.921 | 39 | 516 | 50687 |

**Table I. Estimates of tail index α by the ML method on grouped data**

| Dataset | α | confidence interval | | threshold | k* | n |
|---|---|---|---|---|---|---|
| NL Patval | 0.743 | 0.639 | 0.847 | €3M | 228 | 1046 |
| UK Patval | 0.632 | 0.578 | 0.686 | £650K | 633 | 1368 |

# Conclusions

- Introduction of moving, myopically ‘rational’ firms (self-organizational regime) results in power-law distribution of innovation sizes with tail exponent $\alpha=1$ (as compared to skewed but not fat-tailed distribution for fixed firms)
- This result is relatively insensitive to parameter values
  => self-organized criticality?
- Moving (self-organizing) firms outperform fixed firms
- Collectively more individually ‘rational’ firms innovate faster in general, but there is a saddle for intermediate values of the search radius. Very local or very global search is better than the intermediate case.