Submarkets, Shakeouts and Industry Life-Cycle*

by
Jian Tong, London School of Economics†

Discussion Paper
No EI/26
November 2000

The Toyota Center
Suntory and Toyota International Centers for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel 0207 955 7719

* I am very grateful to John Sutton for his dedicated supervision.

† Correspondence address: Q242, STICERD, LSE, Houghton Street, London WC2A 2AE, UK. Tel: +44-20-7955 6685, Fax: +44-20-7955 6951, E-mail: j.tong@lse.ac.uk

Abstract

Some recent empirical findings suggest that there are intrinsic links between the statistical regularities regarding cohort survival patterns, the persistence of firm turnover, and shakeout during an industry life-cycle. This paper presents a theoretical model which explains these regularities and the links between them. I begin the analysis by treating the market as comprising a number of strategically independent submarkets, so that I can separate the strategic interaction effect at the submarket level and the independence effects which operate across these independent submarkets at the aggregate level. The analysis reveals that within each submarket a 'selection process' in quality competition induces market concentration over time, and this leads to a certain shakeout pattern at the disaggregated level. The study also finds that the dynamics of the emergence of submarkets in a conventionally defined industry plays a crucial role in shaping the aggregate pattern of the industry life-cycle.

Key words: industry life-cycle, industrial growth, submarkets, shake-out.

1 Introduction

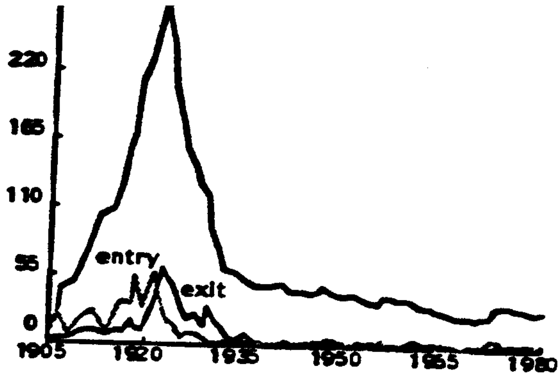
In the past decade the study of industry life-cycle in the firm-growth literature has devoted most of its attention to characterizing and explaining the 'shakeout'[1], i.e., the statistical regularity that the number of producers tends to first rise to a peak and later falls to some lower level in a large number of industries. Recent empirical findings by Horvath et al. (1997) and Klepper (1999) shed light on two other statistical regularities regarding industry life-cycle: (1) the 'turbulence' (firm turnover), i.e., the statistical regularity that the entry-exit process persists throughout an industry life-cycle and that the gross entry and the gross exit are positively correlated; and (2) the cohort survival pattern, i.e., all entry cohorts share a qualitatively similar survival pattern, which displays a significantly higher exit hazard rate at early age than subsequent ages. Furthermore, these might be intrinsically associated with the 'shakeout'. This paper presents a theoretical model to explain these regularities and the links between them.

Panel A of Figure 1 shows a striking example of 'shakeout' and the associated pattern of firm turnover ('turbulence'). Panel B of Figure 1 shows the cohort survival pattern. When the 'turbulence' and the cohort survival pattern are jointly examined, a surprising pattern can be revealed. As first pinpointed by Horvath et al. (1997), despite the fluctuation in entry rates, the timing of exits for different cohorts of entrants is remarkably similar over time: the exit hazard rate is peaked at a very early age of every cohort's life and drops dramatically to low levels for subsequent ages[2]. In this sense, a typical industry life-cycle can be roughly re-described as follows: a miniature shakeout (i.e., an excess entry followed by dramatic early-age exit then followed by gradual subsequent exits) actually happens in the life of each cohort of entrants in a similar way throughout the whole industry life-cycle, and the usually mentioned shakeout is basically an aggregation of these overlapping-cohort miniature shakeout, associated with a gross entry pattern such that the early-stage and late-stage cohorts have small numbers of entrants and the interim cohorts have large numbers of entrants.[3]

[1] See for example, Klepper (1990), Jovanovic and MacDonald (1994), Klepper (1996, 1999).

[2] Note that the vertical axis of Panel B of Figure 1 is in $\log_{10}$ scale, which implies that a seemingly straight line in such a space would actually be as convex as a $\log_{10}$ function if the vertical axis were in a linear scale. Such a convex curve would mean that the earlier the age the higher the exit rate.

[3] It is worth mentioning that such a stylized 'industry life-cycle' is directly related to three of the four well documented statistical regularities, which are about (1) the size-survival-growth relationship and size distribution, (2) age-survival-growth relationship, (3) the shakeout, and (4) the turbulence (or firm turnover). For detailed description on



A.Number of Firms (Tires, US)
horizontal axis: year

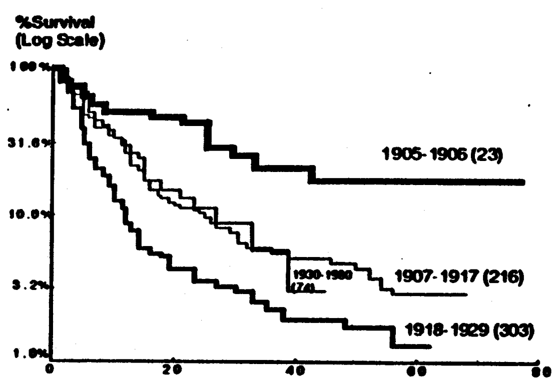


Figure 1: B. Cohort Survival Patern (Tires,US)

horizontal axis: year, vertical axis: survival rate (in log scale)
Source: Klepper 1999

The example of the tires industry, however, is only an extreme case. Dramatic aggregate shakeout proves not a universal phenomenon. According to some research in progress by Steven Klepper, a good example of exception is the laser industry, which has experienced a rather long history of growth but so far has shown no sign of an aggregate shakeout. The divergence regarding the dynamics of aggregate number of producers among industries notwithstanding, one statistical regularity holds firmly among the general run of conventionally-defined (4-digit SIC) industries. That is, there are persistent

these, see Sutton (1997b).

waves of entries over time. The key message conveyed by this simple fact is that independent opportunities keep emerging in a conventionally-defined industry before it matures. This observation matches with the insight described by Sutton (1997a 1997b, 1998), that " most conventionally defined industries exhibit both some strategic interdependence, and some degree of independence across submarkets". If an industry comprises many independent submarkets, then it is natural to see independent opportunities emerge over time, which attract persistent waves of entries. When the notion of independent submarkets *a la* Sutton (1997a 1997b, 1998) is applied to the issue of industrial growth, the logic would suggest that both the pattern of industrial expansion through emergence of independent submarkets, and effect of strategic interaction within each submarket, should leave their fingerprints in the observed pattern of industry life-cycle.

Supposedly, within each submarket the strategic interaction may take the form of price competition with quality choice by producers. Quality, in a broad sense, refers to either the 'consumer perceived quality' or the cost efficiency of a producer. Quality competition takes the form of vertical product differentiation or cost-reducing process innovation, which requires endogenous R&D or advertisement costs. Due to the nonrival and excludable features of 'quality' and the consequent increasing returns to scale, market concentration naturally occurs in industries where quality competition prevails. This scenario is best formalized by Dasgupta and Stiglitz (1980), Sutton (1991, Ch. 3) and Sutton (1998, Ch. 15)[4] in the game-theoretic literature, and is echoed by Klepper (1996, 1999) in the firm-growth literature.

This paper proposes an unusual but plausible extension of the game-theoretic quality competition literature, with an emphasis on independent submarkets, to explain the aforementioned statistical regularities in a typical industry life-cycle. The major scenario to be described is as follows. The uncertainty and informational problems surrounding the beginning of a new submarket tend to impose credit constraints upon fixed expenditures by producers. This in turn restricts the quality competition pressure and leads to the viability of an excessive number of entrants in the early period of the submarket. The miniature shakeout takes place later when the initial credit constraints are removed in light of the resolution of uncertainty and the quality competition pressure is released to terminate the viability of a big fraction of existing producers in the submarket. Since an industry usually contains independent submarkets which take time to emerge and to exhaust the technological potential, the above scenario is repeated over time, and consequently causes 'firm turnover' in the industry. When the emergence of

[4] Sutton (1998) dubs this kind of mechanism the 'escalation mechanism'.

submarkets slows down due to the saturation effect, the gross exit of producers will eventually dominate the gross entry, and an aggregate shakeout takes place.

The remainder of the paper is organized as follows. Section 2 lays out the model. Section 3 characterizes the equilibrium of the model, followed by a simulation of a typical industry life-cycle in Section 4. Some related issues are discussed in Section 5. Section 6 concludes.

2 The model

2.1 'Independent' submarkets

There are S identical consumers in the economy, each of whose utility function has the form of:

$$U = \frac{1}{\Psi(k)} \sum_{i=1}^{k} x_i^{\gamma(k)} + y,$$

where x_i is the consumption of variety i of the 'X' good, k is number of varieties of the 'X' good in the given period, y is the numeraire, standing for all other goods. The difference between this utility function and the usual formulation is that it allows the increase in k to bring some unconventional shocks to the utility function. This feature is embedded in $\Psi(k)$ and $\gamma(k)$ such that $0 < \gamma(k) < 1$, $\frac{\partial \gamma(k)}{\partial k} > 0$ and $\frac{\partial \Psi(k)}{\partial k} > 0$, therefore they capture the idea that the increase of the varieties has a business stealing effect on all existing varieties and it makes all varieties closer substitutes between themselves. The strength of the business stealing effect will depend on the specification[5] of $\Psi(k)$ and $\gamma(k)$. For the sake of simplicity and without loss of generality over the issues to be discussed, it is specified as $\Psi(k) = \gamma(k)$, hence the utility function is specified further to

$$U = \frac{1}{\gamma(k)} \sum_{i=1}^{k} x_i^{\gamma(k)} + y. \tag{1}$$

In any period each consumer maximizes U subject to the budget constraint: $\sum_{i=1}^{k} p_i x_i^{\gamma} + y \leq I$, where I is the total consumption in the given

[5] For example, if it is specified that $\Psi(k) = k$, the utility function will become $U = \frac{1}{k} \sum_{i=1}^{k} x_i^{\gamma(k)} + y$, which implies a very strong business stealing effect.

period[6]. The first order condition of this maximization program is:

$$x_i^{\gamma-1} = p_i,$$

which implies that the demand function of variety i is

$$X_i = S\left(\frac{1}{p_i}\right)^{\frac{1}{1-\gamma}}, \tag{2}$$

where $X_i \equiv Sx_i$.

The price elasticity of demand is constant in each period as follows,

$$\zeta \equiv -\frac{\partial \ln X_i}{\partial \ln p_i} = \frac{1}{1-\gamma}, \tag{3}$$

for which $\frac{\partial \gamma(k)}{\partial k} > 0$ implies $\frac{\partial \zeta}{\partial k} > 0$ (elasticity enhancing effect).

The above specific utility function implies that the business stealing effect notwithstanding, the demand over each variety is only dependent on its own price in any given period. In other words, the submarkets of the 'X' industry are strategically independent of each other.

2.2 Industrial growth via the emergence of submarkets

The above specific utility function also implies that the growth of the industry is through the emergence of new submarkets. It is further assumed in this paper that the number of submarkets in period t, i.e., $k(t)$, follows a generalized Logistic diffusion curve:

$$\begin{cases} \dot{k} = ak^{\theta}(b-k)^{\lambda} \\ k(t) = 0 \text{ for } t < 0 \\ k(0) = 1 \end{cases}, \tag{4}$$

where b ($b \geq 1$) is the saturation number of independent submarkets within the industry. This law of motion captures the feature that growth of the number of submarkets is dependent on the existing number and the potential which hasn't been fulfilled. This feature can be demonstrated easily with the example: $\theta = \lambda = 1$, which simplifies the law of motion to:

$$\dot{k} = ak(b-k),$$

[6] Each consumer's intertemporal optimization program is completely trivialized by the specification of a quasi-linear utility function and unity discount factor, and the implicit assumption that all 'X' goods are unstorable.

to which the closed-form solution is[7]:

$$k = \frac{be^{abt}}{c + e^{abt}},$$

where c is a constant. The implied rate of emergence of submarkets is:

$$\dot{k} = \frac{ab^2ce^{abt}}{\left(c + e^{abt}\right)^2}.$$

Figures 2 and 3 show some general features of the generalized Logistic diffusion curve: (1) it is initially convex up to some point, then (2) it becomes concave, and finally (3) it becomes flat. Accordingly, the growth rate initially increases up to a peak, then declines, finally converges to zero.

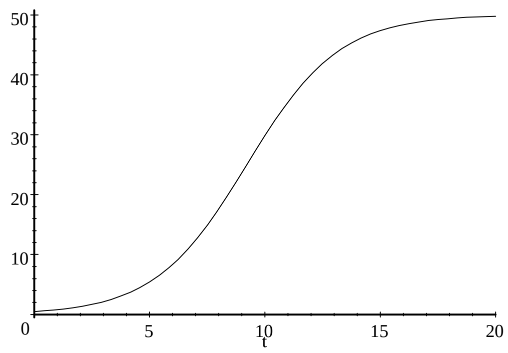


Figure 2: $k(t)$

The second order derivative of $k(t)$ is as follows,

$$\ddot{k} = a(\theta + \lambda) k^{\theta-1} (b - k)^{\lambda-1} \left(\frac{\theta b}{\theta + \lambda} - k \right) \gtreqless 0 \text{ when } k \lesseqgtr \frac{\theta b}{\theta + \lambda},$$

which reveals that the peak of growth rate is located at the point such that $k = \frac{\theta b}{\theta+\lambda}$.

2.3 Repeated 3-stage game in each submarket

Whenever a new independent submarket emerges, the following game starts to be played.

[7]Note that $\frac{1}{k}\frac{dk}{dt} + \frac{1}{b-k}\frac{dk}{dt} = ab$ and $\frac{d}{dt} \ln \frac{k}{(b-k)} = ab$, which imply the displayed general solution.

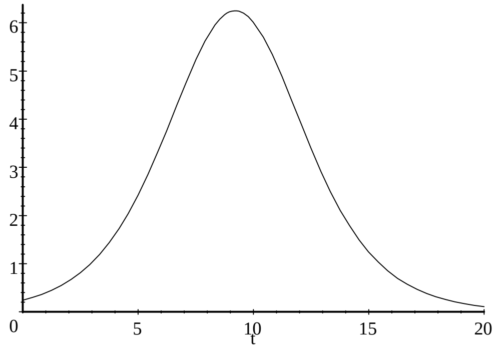


Figure 3: $\dot{k}(t)$

Phase One (credit-constrained phase):

Stage one: entry decision when there is credit constraint $\rightarrow N_1$,
Stage two: quality (or productivity) choice,
Stage three: quantity competition (Cournot).

Phase Two (credit-unconstrained phase):

Stage one: entry decision when there is no credit constraint$\rightarrow N_2(t)$,
Stage two: quality (or productivity) choice,
Stage three: quantity competition (Cournot).

The Phase One game is played only in one period, and the Phase Two game is repeated in each subsequent period. The rationale of this game is that the first period, which is the beginning of the submarket, is marked by the uncertainty about each producer's capability of handling the technology of improving quality, therefore no external finance is involved in the investments in quality. As a result each producer's quality choice is restricted by a credit constraint. At the end of the first period, the uncertainty is resolved. Only those players who have proven high capabilities enter Phase Two, whose credit constraints are removed. In other words, after one period in the history of each submarket, a wave of selective financial shocks will happen to the efficient players and transform them into credit-un-constrained players. This consequently changes the phase of the game. It is further assumed that only the players who have successful track records in the previous period are eligible for re-entry in any period of phase two.

The above description is the idea which motivates our modelling. However for the sake of simplicity, we will not explicitly model the efficiency-based financial selection process. Instead we will assume that some random selection takes place among the symmetric players, which selects a particular equilibrium among all the possible symmetric equilibrium outcomes.[8]

2.4 The games of quality competition with or without credit constraints

It is assumed that the quality achieved by each player in any period is not carried over to the next period due to depreciation. Under this assumption, except for a shift of parameter, each period of the repeated game almost becomes an isolated game, the equilibrium of which can be characterized without reference to what happens elsewhere. Therefore the subgame perfect Nash equilibrium is a sufficient solution concept for characterizing the game formulated in this paper.

The three-stage game in each submarket during each period then can be solved by backward induction. The third stage subgame is always a Cournot game of quantity competition, for which a Cournot-Nash equilibrium exists and is unique up to a given set of quality levels of all incumbents. The subgame Nash equilibrium determines for each producer a reduced form profit function as follows[9]:

$$\pi_i\left(u_i \mid u_{-i}\right) = S\left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}} \frac{1}{1-\gamma}\left(1-\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)^2, \qquad (5)$$

where u_i is the quality level of producer i, u_{-i} is a $(N-1)$-tuple of quality levels of other producers except producer i, for $i = 1, 2, \cdots, N$; N is the number of producers active in the submarket and S is the population of consumers in the economy.

Interested readers can find the derivation of this reduced form profit function in the Appendix. The key feature of it is that a producer's profit increases with its relative quality level against its rivals', i.e., $\sum_{j=1}^{N}\frac{u_i}{u_j}$. This function tells how the strategic environment responds to vertical differentiations of producers. It captures the Darwinian selection pressure embedded in the environment constituted by customers and rivals. This environment provides producers incentives to outperform their rivals in R&D and quality.

[8] In a more complicated version of the current paper, a mechanism which may generate this kind of selective financial shocks is explicitly discussed.

[9] This formulation is an extension of the one developed in Appendix 15.1 of Sutton (1998). It allows a general $\gamma \in [0, 1)$ while the Sutton (1998) formulation deals with a special case: $\gamma = 0$.

When they vie in quality, each of them has to bear the fixed cost, which is a function of quality level as shown below:

$$F(u_i) = \begin{cases} \mu u_i^\beta & \text{in credit-un-constrained phase} \\ \begin{cases} \mu u_i^\beta & \text{if } \mu u_i^\beta \le \delta \\ \infty & \text{if } \mu u_i^\beta > \delta \end{cases} & \text{in credit-constrained phase} \end{cases}, \tag{6}$$

where δ is the credit limit in the first period. The above fixed cost function is depicted in Figure 4, which shows that in the credit constrained phase, the effect of the credit constraint (δ) is equivalent to putting an upper bound to the quality level.

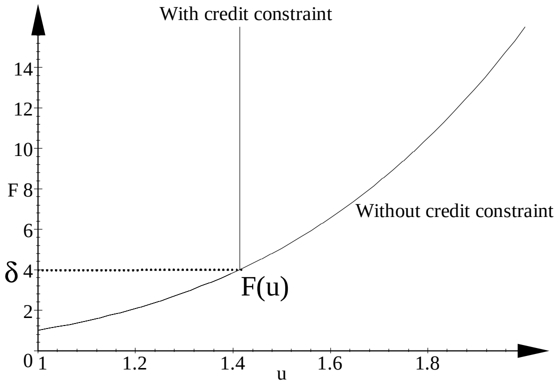


Figure 4: Fixed cost functions

Therefore each producer's objective in second stage subgame is to maximize its net profit by choosing own quality level given others' quality levels,

$$\max_{u_i} \left\{ S \left(\frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{1}{u_j}} \right)^{\frac{\gamma}{1-\gamma}} \left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)^2 - F(u_i) \right\}. \tag{7}$$

If we take the above payoff function as given, then the game in each period is merely a simple game of quality competition with free entry.

3 Equilibrium

3.1 Market structure in credit-un-constrained phase

We proceed to characterize the equilibrium, starting with Phase Two, when credit constraints have been removed for remaining players. The focus is on a

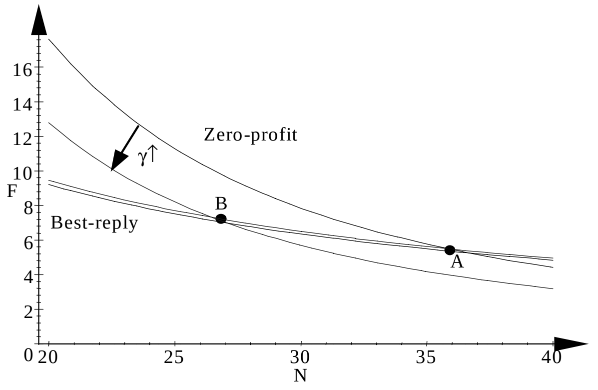


Figure 5: Equiliria without credit constraints

symmetric subgame perfect Nash equilibrium, where each producer equates the marginal benefit of increasing quality and the marginal cost. The implied best-reply function for each producer in a symmetric outcome is

$$F = \mu u^{\beta} = \frac{S}{\beta}\left(\left(1-\frac{1-\gamma}{N_2}\right)u\right)^{\frac{\gamma}{1-\gamma}}\left(2\left(N_2-1\right)\frac{N_2-1+\gamma}{N_2^3}+\frac{\gamma}{N_2^3}\right), \quad (8)$$

where N_2 is the number of producers in a Phase Two equilibrium.

Free entry into the submarket results in the following zero profit condition in a symmetric outcome:

$$F = \mu u^{\beta} = S\frac{1-\gamma}{N_2^2}\left(\left(1-\frac{1-\gamma}{N_2}\right)u\right)^{\frac{\gamma}{1-\gamma}}. \quad (9)$$

The above two conditions pin down the equilibrium submarket structure and the equilibrium endogenous fixed cost level as shown in Figure 5. The graph also indicates that the equilibrium without credit constraints depends on parameter γ: when γ increases the submarket structure becomes more concentrated (point **B** vs. point **A**), i.e., N decreases.[10]

Proposition 1 [11]*In a symmetric Phase Two equilibrium, the number of producers is* $N_2 = n(\beta,\gamma) \equiv \frac{n_0+\sqrt{n_0^2-2(2-\gamma)}}{2}$, *where* $n_0 \equiv 1+\frac{(\beta+2)(1-\gamma)}{2}$. *Furthermore* N_2 *is decreasing in* γ.

Proof. See Appendix A.2.

[10] An increase in γ shifts the Zero-profit curve downward and to the left, but shifts the Best-reply curve upward and to the right. The net effect however, is dominated by the former. In fact, as illustrated in the graph the second effect is quantitatively very small compared to the first, so it can be ignored for the sake of simplicity.

[11] The result is consistent with that found in Dasgupta-Stiglitz (1980).

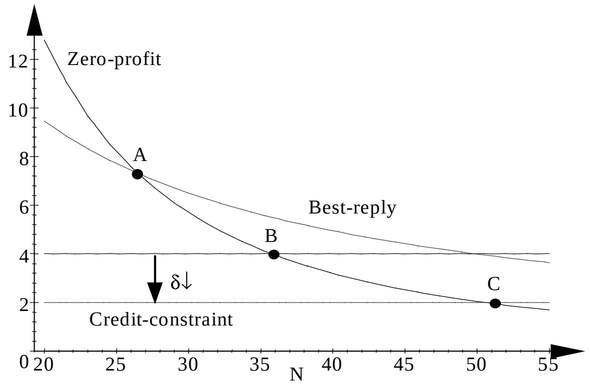


Figure 6: Equilibria with and without credit constraints

3.2 Market structure in credit-constrained phase

It is assumed that with the credit constraints, the credit-un-constrained equilibrium fixed cost level is not feasible, therefore the credit constraints must be binding. So the following binding credit constraint should replace the best reply condition in determining the equilibrium fixed cost level:

$$F_c = \mu u_c^{\beta} = \delta. \tag{10}$$

Accordingly, the zero profit condition should be modified to:

$$F_c = \mu u_c^{\beta} = S\frac{1-\gamma}{N_1^2}\left(\left(1-\frac{1-\gamma}{N_1}\right)u_c\right)^{\frac{\gamma}{1-\gamma}}, \tag{11}$$

where N_1 is the number of producers in the Phase One (credit-constrained) equilibrium.

The above two conditions pin down the equilibrium submarket structure and the equilibrium endogenous fixed cost level as shown in Figure 6. It can be seen from the graph that the submarket is more fragmented when there is credit constraint (point **B**) than when there isn't (point **A**), yet the tighter the credit constraints the more fragmented the submarket (point **C** vs. point **B**).

Proposition 2 *Under the assumption: $\beta > \frac{\gamma}{1-\gamma}$, the tighter the binding credit constraints are, the more fragmented the submarket is. Also, there is a lower bound to N_1 such that $N_1 > n(\beta, \gamma)$, i.e., when the credit constraints are strictly binding, the submarket is more fragmented than if there were no credit constraints.*

Proof. See Appendix A.3.

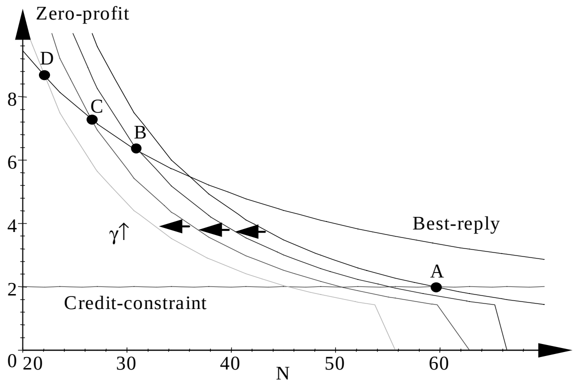


Figure 7: Evolution of submarket structure

3.3 Miniature shakeout and cohort survival pattern

Since all submarkets which emerge in the same period are similar, any cohort dynamics is merely a reflection of the dynamics of a representative submarket which belongs to that cohort. So a good point of departure to understand cohort dynamics is to look at what kind of shocks a typical submarket experiences over time.

As the industry grows, it becomes 'tighter' in the sense that it is filled with more submarkets. Consequently, the price elasticity of demand in each submarket increases. This insight is captured by $\frac{\partial \gamma_t}{\partial k} > 0$, which implies:

$$\frac{d\gamma_t}{dt} = \frac{\partial \gamma_t}{\partial k}\frac{dk}{dt} > 0. \tag{12}$$

For a representative submarket, this means a series of shocks to the price elasticity of demand over time. By Proposition 1, the effects of these shocks should make the submarket more and more concentrated, which implies persistent exit of producers over time. On top this kind of exit pressure, the selective financial shocks which occur in the first period of the representative submarket may cause a more dramatic wave of exit of incumbents.

The evolution of the of structure of a typical submarket is demonstrated by Figure 7, where point **A** stands for the initial credit-constrained equilibrium and points **B**, **C** and **D** stand for the subsequent credit-un-constrained equilibria. The jump from **A** to **B** is caused by two waves of shocks: the selective financial shocks and the shock to demand elasticity. Subsequent shifts: from **B** to **C**, from **C** to **D** and so on are only due to shocks to demand elasticity[12], therefore are less dramatic. The typical cohort survival

[12] The increases of γ should shift the best-reply curve upward and to the left. Since the

pattern then is in full accordance with the evolution of submarket structure and is described by the following

Proposition 3 *Each cohort experiences a sequence of shocks such that after one period, there is a wave of selective financial shocks, and in all periods there are gamma shocks which shift the parameter γ upward. Consequently, each cohort experiences a miniature shakeout after one period, and continuous decline of number of producers in subsequent periods. If the binding credit constraints are sufficiently tight then each entry cohort has a higher first period hazard rate than subsequent periods.*

3.4 Aggregate shakeout

To an observer who can only see what is going on in the whole industry rather than in each entry cohort, let alone each submarket, what she observes is the aggregate pattern. The aggregate pattern depends, for any given cohort survival pattern, on the industry-specific characteristics regarding growth, which can be captured by a set of parameters such as (a, b, θ, λ), embedded in the law of motion of industrial growth: $\frac{dk}{dt} = ak^{\theta}(b-k)^{\lambda}$.

For example, if in a conventionally-defined industry which consists of a large number of strategically independent submarkets and the link between each submarket and the rest of the industry is very weak, then b should be sufficiently large, $\theta \to 0$ and $\lambda \to 0$. Consequently the observer would not see a dramatic aggregate shakeout at all, instead, she would be observing a very slow but steady build-up of number of producers during a long period in that industry. This pattern can arise due to the extreme specification of the generalized logistic curve:

$$\frac{dk}{dt} = ak^{\theta}(b-k)^{\lambda} \to a.$$

But more likely, she would be able to see some kind of aggregate shakeout when the industrial growth characteristics are not so extreme. Generally, the larger θ and λ in a industry, the more dramatic the aggregate shakeout. In the next section, we calibrate the parameters of the model and simulate the process of an industry life-cycle. Two different cases will be distinguished with respect to the aggregate pattern.

---

effects of these changes are dominated by the effects of the changes to zero-profit curve, therefore are ignored in the graph, without causing non-trivial bias.

4 Simulations of industry life-cycle

4.1 Case 1: with dramatic aggregate shakeout

The specification of the k-γ relationship is as follows, $\gamma = 1 - \epsilon \exp(\omega(-k+1))$. Since the major interest of current study is to understand the working of some economic mechanisms rather than achieving high level of descriptive realism, the current calibration of the model for the simulation hasn't yet been carefully guided by real world data. The specification of the parameters are shown in the following table.

Parameter:	a	b	θ	λ	β	δ	S	μ	ϵ	ω
Value:	0.0004	50	1.8	1.5	200	10	200000	1	0.9	0.05

The rough simulation nevertheless confirms the capability of the model to predict the typical patterns of aggregate shakeout, persistent turbulence and cohort dynamics. Figure 8 depicts the simulated aggregate shakeout and the correlated gross entry and gross exit over time. This graph resembles the real world picture presented in Panel A of Figure 1. The predicted correlation between gross entry and gross exit over time is 0.76, which is comparable with an estimated number from another source[13]. Figure 9 further breaks the aggregate shakeout pattern down to separated cohort dynamics. It suggests that the dramatic rise and fall of the total number of producers around the peak are largely due to impacts of a very few big cohorts (9, 10 and 11), which feature "mass entry waves followed by mass exit waves". The pattern of 'miniature shakeout' is demonstrated by Figure 10, which shows that the early-age exit hazard is significantly higher than subsequent ages in each cohort. The consequent typical cohort survival pattern then is presented in Figure 11, which shows the average survival rates over age. The simulated pattern qualitatively resembles the real life pattern presented in Panel B of Figure 1.

[13] Empirically, within any one country, a strong correlation is found to exist between entry and exit rates by industry. For example, Paul Geroski (1991) reports a correlation coefficient of 0.796 for a sample of 95 industries in the U.K. in 1987.

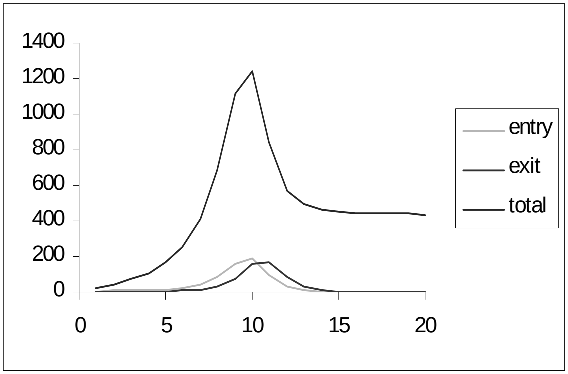


Figure 8: vertical axis: number of producers, horizontal axis: time (simulation 1)

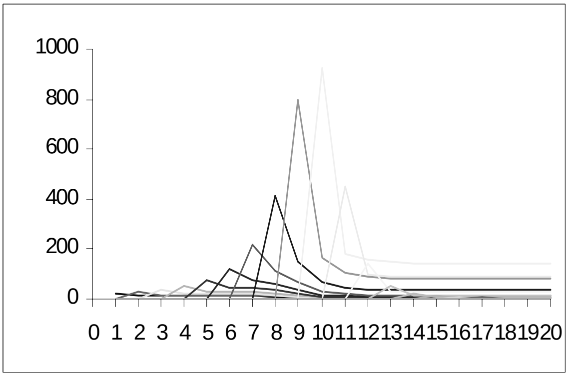


Figure 9: Cohort dynamics (simulation 1), vertical axis: number of producers, horizontal axis: time

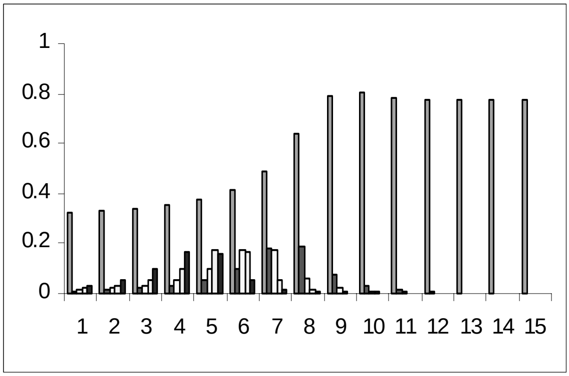


Figure 10: Cohort exit rates: age 2-5 (simulation 1), horizontal axis: cohort

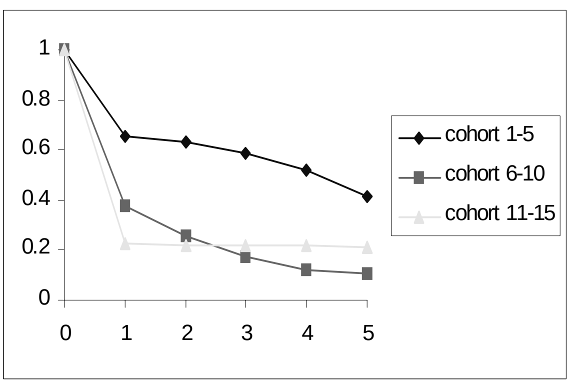


Figure 11: Average cohort survival pattern (simulation 1), horizontal axis: age

4.2 Case 2: without aggregate shakeout

Now we proceed to verify the point that if an industry consists of a large number of submarkets which have very weak links with the rest of the industry, i.e., b is sufficiently large and θ and λ are sufficiently small, then dramatic aggregate shakeout should not take place. The following specification of parameters embodies these features.

Parameter:	a	b	θ	λ	β	δ	S	μ	ϵ	ω
Value:	0.4	50	0.01	0.01	200	10	150000	1	0.9	0.0005

The simulation confirms the conjecture. Figure 12 shows a simulated industry life-cycle without an aggregate shakeout. It is worth mentioning that this result does not rely on any qualitative difference in cohort survival pattern. Figure 13-15 show that the simulated cohort survival pattern in this new specification is qualitatively similar to the previous one.

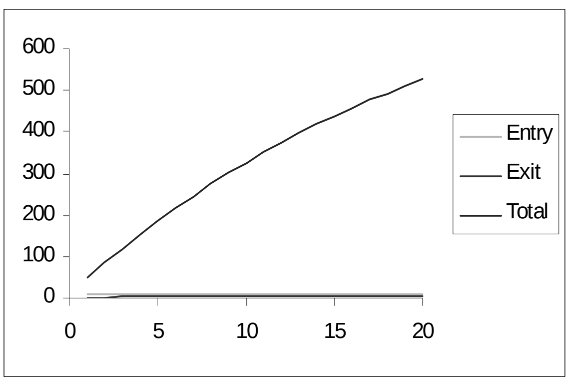


Figure 12: vertical axis: number of producers, horizontal axis: time (simulation 2)

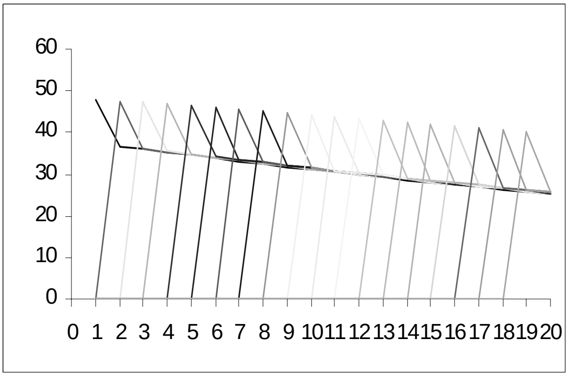


Figure 13: Cohort dynamics (simulation 2), vertical axis: number of producers, horizontal axis: time

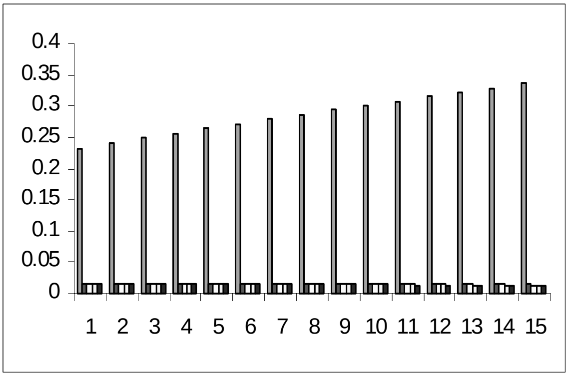


Figure 14: Cohort exit rates: age 2-5 (simulation 2), horizontal axis: cohort

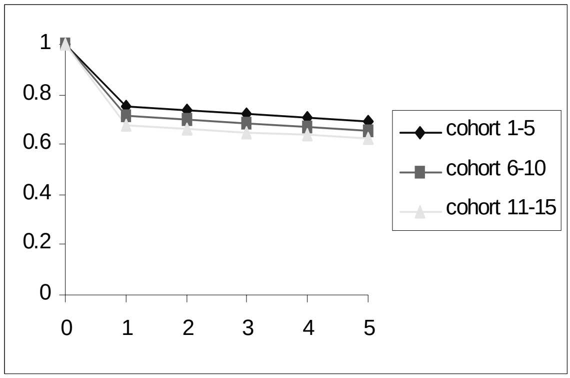


Figure 15: Average cohort survival pattern (simulation 2), horizontal axis: age

5 Discussion

5.1 Price and productivity dynamics

In every theoretical model, there is a trade-off between analytical simplicity and descriptive realism. To keep the analysis simple, two of the stylized facts about the dynamics of price and productivity, which have attracted research interest in the literature, haven't been addressed in this study. To model the effects of knowledge accumulation embedded in the increase of the number of varieties in the industry, the model could be easily extended to allow for the fixed cost function in each submarket to shift downward over time in connection to this kind of knowledge accumulation. Then the model would be able to predict the pattern that prices decline over time with the increase of productivity (or quality).

5.2 Multi-product firms

A major caveat is in order. The empirical description of an industry life-cycle usually uses firm as the unit of measurement. The unit of count employed in the current analysis is business. In the real world it is quite common that some firms are multi-product firms, which can be regarded as a bundle of many segments operating in different submarkets. Since the actions of entry or exit regarding a particular submarket by a multi-product firm may be obscured in the data by its presence in other submarket(s), the match between the simulated pattern based on the present model and the observed pattern must be distorted by some bias. The bias is caused by an implicit assumption that each firm consists merely of one business segment. The stake at issue is: will this conceivable bias disqualify the results of the analysis?

The answer will depend on the relation between firms and submarket players. It is very likely that a reasonably stable proportion of firms are solo-submarket players, so their entry and exit actions which can be observed are quite representative for the total entry and exit actions regarding submarkets. Therefore it is unlikely that the conceived bias could disqualify the main findings of the study. The caveat then is that if this model is to be extended to the study of firm size distribution, the aforementioned implicit assumption on composition of a firm must be replaced by a more realistic one for that task.

5.3 Multi-equilibria and indeterminacy

It is noticed that due to the existence of identical multi-equilibria in each submarket there is intrinsical indeterminacy about which players should survive in which submarkets, this naturally leaves room for randomness in the outcome, though the setup of the model is deterministic. This feature makes the model consistent with the key notion in the firm growth literature that firms grow randomly.

5.4 Industry characteristics and life-cycle

In the previous section we have simulated two extreme cases of industry life cycles, which have indicated a possible wide range of difference between industries. A interesting question is: Can this kind of difference be explained against difference in industry characteristics? The point which can be made in light of the above analysis is, should investigation proceed along this line, some characteristics regarding the dynamic feature as well as financial feature of an industry deserve particular attention. Such candidates should include the diversity of the industry, reflecting the number of independent submarkets, the expansion features such as growth rate and the trends of growth rate, dependence on external finance, and degree of credit constraint in respect to investments in intangible assets.

6 Concluding remarks

This study, while following the tradition of the growth-of-firms literature in examining the 'selection process' in product market to gain insight of the industry life-cycle, moves a first step in a new dimension of understanding this object. By addressing the issue of the emergence of submarkets, it segregates two levels of analysis: the submarket level and the aggregated industry level.

At the submarket level, a selection process is suggested to play the primary role in generating the miniature shakeout pattern. A novel feature of our analysis is that we characterize the roles of imposing and removing credit constraints upon the quality competition in the changes of competitive structures. At the aggregated industry level, we notice that a generalized Logistic diffusion curve may fit in the job of describing the dynamics of emergence of submarkets.

This analysis rests on the notion of 'independent submarkets' and demonstrates a new application of it in providing a useful benchmark to the study of market structure. In doing so, the analysis casts some new light on the potential of the program pioneered by John Sutton (1997a, 1997b), which is to build game-theoretic models of markets that contain independent submarkets in an aim to explaining the well-documented statistical regularities over the evolution of market structure. In this paper, the 'shakeout' and the 'turbulence' have been explained to a substantial extent. With a view to the fact that a conventionally-defined 'firm' in the empirical studies can be a bundle of segments active in several independent submarkets, it would be possible to extend the model to predict certain impacts of age and size on survival. The postulated simple rationales can be that bigger firms have bigger portfolios of independent segments and therefore have lower overall exit hazard rate, that for any given firm size older firms are more likely to have portfolios which include smaller proportion of high-risk pre-shakeout segments and therefore have lower overall exit hazard rate.

A Appendix

A.1 Derivation of the reduced form profit function

The subscript can be removed when a representative submarket in an arbitrary period after its emergence is addressed. So the demand function in eq. (2) becomes:

$$X = S\left(\frac{1}{p}\right)^{\frac{1}{1-\gamma}},$$

where[14] X is the submarket's output, $X = \sum_{j=1}^{N} x^j$, x^j is the output of producer j and N is the number of producers. It can be rewritten as

$$p = \left(\frac{S}{\sum_{j=1}^{N} x^j} \right)^{1-\gamma}.$$

Consider the Cournot Competition in quantity: for given vector (x^j), $j \neq i$, producer i chooses x^i to maximize her profit:

$$\max_{x^i} \left(\left(\frac{S}{\sum_{j=1}^{N} x_j} \right)^{1-\gamma} - \frac{1}{u_i} \right) x^i,$$

where $\frac{1}{u_i}$ is the cost of producing one unit of effective quantity. u_i is the 'quality' (or productivity) index. The FOC of the program is:

$$-(1-\gamma) \frac{\left(\frac{S}{\sum_{j=1}^{N} x^j} \right)^{1-\gamma}}{\sum_{j=1}^{N} x^j} x^i + \left(\frac{S}{\sum_{j=1}^{N} x^j} \right)^{1-\gamma} - \frac{1}{u_i} = 0.$$

Substituting $\left(\frac{S}{\sum_{j=1}^{N} x^j} \right)^{1-\gamma}$ with p and solving for x^i gives the following equation:

$$x^i = \frac{\left(p - \frac{1}{u_i} \right) X}{(1-\gamma)\, p}.$$

The submarket share of producer i then is

$$\frac{x^i}{X} = \frac{1}{1-\gamma} \frac{p - \frac{1}{u_i}}{p}.$$

Summing up the submarket shares of all producers to eliminate x^i and X: $1 = \sum_{i=1}^{N} \frac{x^i}{X} = \frac{1}{1-\gamma} \frac{Np - \sum_{i=1}^{N} \frac{1}{u_i}}{p}$ and solving for p leads to

$$p = \frac{\sum_{j=1}^{N} \frac{1}{u_j}}{N - 1 + \gamma}.$$

By the above equations the submarket share of producer i can be solved as $\frac{x^i}{X} = \frac{1}{1-\gamma} \left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N} \frac{u_i}{u_j}} \right)$.

[14] If X is interpreted as effective quantity, then this demand function can be applied to both homogeneous good or vertically differentiated good.

The mark-up ratio of producer i is $\frac{p-\frac{1}{u_i}}{p} = \left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)$.

The submarket total output in equilibrium then is

$$X = S\left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{1}{1-\gamma}}.$$

The total revenue of the submarket turns out to be

$$pX = S\left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}}.$$

Finally the reduced form profit function of producer i is derived as

$$\begin{aligned}\pi_i\left(u_i \mid u_{-i}\right) &= \left(p - \frac{1}{u_i}\right)x^i = S\left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)^2 \frac{1}{1-\gamma}\left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}} \\ &= \underbrace{S\left(\frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{1}{u_j}}\right)^{\frac{\gamma}{1-\gamma}}}_{\text{total sales}} \underbrace{\frac{1}{1-\gamma}\left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)}_{\text{submarket share of producer } i} \underbrace{\left(1 - \frac{N-1+\gamma}{\sum_{j=1}^{N}\frac{u_i}{u_j}}\right)}_{\text{mark-up ratio of producer } i}.\end{aligned}$$

A.2 Proof of Proposition 1

By eq. (8) and eq. (9) the following quadratic equation can be derived: $N_2^2 - n_0 N_2 + \frac{2-\gamma}{2} = 0$, the largest positive solution of which is $N_2 = \frac{n_0 + \sqrt{n_0^2 - 2(2-\gamma)}}{2}$. It is not difficult to verify that $\frac{\partial N_2}{\partial \gamma} < 0$.

A.3 Proof of Proposition 2

Combining eq. (10) and eq. (11) leads to the following equation: $S\frac{1-\gamma}{N_1^2}\left(\left(1 - \frac{1-\gamma}{N_1}\right)\left(\frac{\delta}{\mu}\right)^{\frac{1}{\beta}}\right)^{\frac{\gamma}{1-\gamma}} = \delta$. Taking the logarithm transformation of both sides of the above equation and manipulating it can lead to: $\ln\left(S\left(1-\gamma\right)\right) - \frac{\gamma}{1-\gamma}\frac{1}{\beta}\ln\mu - 2\ln N_1 + \frac{\gamma}{1-\gamma}\ln\left(1 - \frac{1-\gamma}{N_1}\right) = \left(1 - \frac{\gamma}{1-\gamma}\frac{1}{\beta}\right)\ln\delta$. Differentiating both sides w.r.t. δ and reorganizing the equation reveals the following derivative: $\frac{\partial N_1}{\partial \delta} = -\frac{\left(1-\frac{\gamma}{1-\gamma}\frac{1}{\beta}\right)N_1(N_1-1+\gamma)}{2N_1-2+\gamma}$. It is easy to see that $\frac{\partial N_1}{\partial \delta} < 0$ when $\beta > \frac{\gamma}{1-\gamma}$.

References

[1] Dasgupta, Partha and Joseph Stiglitz (1980), "Industrial Structure and Innovation Activity", The Economic Journal, 90, 266-293.

[2] Hopenhayn, Hugo A. (1993), "The Shakeout", Economics Working Paper 33, Universitat Pompeu Fabra.

[3] Horvath, Michael, Fabiano Schivardi and Michael Woywode (1997), "On Industry Life-Cycle: Delay, Entry, and shake-out in Beer Brewing", Mimeo, Stanford University.

[4] Jovanovic, Boyan (1979), "Selection and the Evolution of Industry", Econometrica, vol. 50, no. 3, pp. 327-38.

[5] Jovanovic, Boyan and Glenn M. MacDonald (1994), "The Life Cycle of a Competitive Industry", Journal of Political Economy, vol. 102, no. 2, 322-47.

[6] Klepper, Steven and Elizabeth Graddy (1990), "The Evolution of New Industries and the Determinants of Market Structure", Rand Journal of Economics, 21(1), pp. 27-44.

[7] Klepper, Steven (1996), "Entry, Exit, Growth, and Innovation Over the Product Life Cycle", American Economic Review, 86(3), pp. 562-83.

[8] Klepper, Steven (1999), "Firm Survival and the Evolution of Oligopoly", Mimeo, Carnegie Mellon University.

[9] Sutton, John (1991), Sunk Cost and Market Structure, Cambridge, Massachusetts: The MIT Press.

[10] Sutton, John (1997a), "Game Theoretic Models of Market Structure." Advances in Economics and Econometrics, edited by D. Kreps and K. Wallis. (Proceedings of the World Congress of the Econometric Society, Tokyo, 1995.) Cambridge: Cambridge University Press, pp. 66-86.

[11] Sutton, John (1997b), "Gibrat's Legacy", Journal of Economic Literature, vol. XXXV (March), pp. 40-59.

[12] Sutton, John (1998), Technology and Market Structure, Cambridge, Mass: The MIT Press.