Expansion Economies:
The Growth of the Firm through its Globalization

Fernando Buendía
School of Economic and Administrative Sciences
Universidad Panamericana-Campus Guadalajara
Prolongación Calzada Circunvalación Poniente # 49
Ciudad Granja
Zapopan, Jalisco, Mexico, 45010
Phone: +52 (33) 13-18-22-18
Fax: +52 (33) 13-68-22-01
E-mail: fernando.buendia@up.edu.mx

Paper

for
The 30th International Conference of the System Dynamics Society
St. Gallen, Switzerland
July 22 – 26, 2012

Abstract

Using the tools of system dynamics, this paper formalizes the theory of expansion economies to explain the globalization of the firms. In order to exploit economies of expansion, manufacturing firms tend to expand their economic activities to different locations, regions and countries. As firms' expansion process is an increasing return mechanism, system dynamics and urn theory can explain path dependence and self-organizing size distribution of global firms.

1. Introduction

Agglomeration economies are fundamental to understanding the increasing divergence in the allocation of economic activity among countries and the economic growth of the countries where those areas are located. Even though economic geography has gained a relevant place in economic analysis and a large influence in the theory of international trade, its findings cannot be generalized to all kind of firms and industries. It is true that in certain cases centripetal forces caused by agglomeration economies can compel firms to cluster in a given geographical area and become localized. In other cases, however, firms have incentives to become *global*. Furthermore, that firms expand their economic activities to different geographical areas is one of their most remarkable features. For instance, traditional firms such as Ford, General Motors, and Coca-Cola, and recent enterprises, such as Wal-Mart and Starbucks, have extended their operations in a large number of locations. Using the notion of expansion economies (Buendía, 2006) and the tools of systems dynamics theory, in this paper I develop a model of the globalization of the firm and show that increasing returns to the growth of the firm due to expansion economies produce self-organizing industrial structures. Expansion economies, therefore, are a fundamental theoretical foundation to revise the theory of industrial clusters and international trade theory. This paper has four additional sections. Section two discusses the theoretical

background of the theory of the growth of the firm. In the third section, the theory of expansion economies is explained. In section four, using the tools of system dynamics, the notion of expansion economies is formalized. Section five discusses how urn theory can be used to formalize systems where expansion economies play an important role and cause slated distribution of sizes of firms.

2. The Foundations for an Economic Theory of the Expansion of the Firm[1]

The growth of the firm through its globalization has received little attention in standard economic analysis. Particularly, economic geography has shown that economic activity is concentrated geographically in few regions —leaving other regions relatively undeveloped— and, with this, it has explicitly assumed that firms tend to be little globalized. Although the model of dispersion economies developed by Polenke (2003) has shed light on the benefits of dividing up a firm to reduce pollution, it seems accurate to say that expansion economies still has a limited place in economic theory. When the evolution of the modern corporation is considered, this lack of interest becomes rather surprising because the geographic expansion of firms —as opposed to its concentration in a large-scale production unit in specific geographical area— is perhaps one of the most common strategies many firms have adopted to grow.

The relevance of the growth of the firm through the expansion of its economic activities to other locations becomes evident when it is compared with the traditional theories of the firm. Coase's (1988) transaction costs approach and Williamson's (1975) contracts perspective are of obvious importance to understand why firms exist and how they can reduce transactions costs. But they say little about how the firm grows. With his seminal article *The Economies of Scale,* Stigler (1958) laid the foundations of the theory of the growth of the firm. His argument is that the more rapid the rate to which a firm loses its share of the industry's output (or capacity), the higher is its private cost of production relative to the cost of production of firms of the most efficient size. However, this is more a theory of the optimum size of plant (or manufacturing capacity of the firm) than a theory of the firm of optimum size[2]. This distinction is important, for many economists have thought of scale economies as those economies stemming not only from plants of optimum size, but also from the vertically integrated firms. However, the economies that a vertically integrated firm obtains are radically different from Stigler's original interpretation of scale economies.

As a matter of fact, the earliest discussion of the properly combination of all the productive services that are part of the normal structure of the firm of best possible size has to be attributed to Chandler (1966, 1977, and 1990). Chandler's main intellectual contribution was to recognize that, in order to achieve the lower unit costs, firms had to do a lot more than simply build large plants. They had to be able to maintain a high rate of throughput through their factories —that is, to keep their plants operating consistently at high levels of capacity utilization. In order to maintain a high rate of throughput, firms had

[1] This and the following section draw heavily from Buendía (2006).

[2] Although Stigler (1958) sets out his paper by saying that the theory of economies of scale is the theory of the relationship between the scale and the use of a properly chosen combination of all productive services and the rate of output of the enterprise, his analysis just focuses on the manufacturing services.

to insure that shortfalls in supply did not disrupt their production processes and that output did not pile up in their warehouses unsold. The solution, as Chandler saw it, was for firms to bring these activities under their direct control by integrating backward into raw-material production and forward into distribution, and by building a managerial hierarchy capable of coordinating smoothly the flow of inputs and outputs from raw material to final sale. Therefore, through his historical theory of large business, Chandler has provided empirical evidence of the existence of what theoretically can be called economies of integration.

Nevertheless, large firms could exploit not only economies of scale and economies of integration, but also economies of scope. According to Chandler, large firms can reap economies of scope by investing large quantities of financial resources in research and development, which allows them diversify their operations into other industries. Chandler claimed that firms that reaped scale economies, integration economies and scope economies improved upon the workings of the market, captured the resulting efficiency gains, obtained enormous competitive advantages, and over time brought under their managerial authority larger and larger portions of the economy. The only firms that could compete with them head to head, he argued, were those that completely duplicated their vertically integrated structures and managerial hierarchies. Because relatively few firms could raise the enormous amounts of capital required, these kinds of industries quickly took on oligopolistic structures.

When The Visible Hand was first published in 1977, Chandler's synthesis represented an extraordinary achievement. It provided a respectable alternative to the robber-baron view of big business that still figures prominently in the industrial economics literature. It also offered business historians for the first time a framework that made sense of the many (often antiquarian) histories of individual firms and industries that to that point largely constituted the field. Most significantly, it focused its attention on the central economic problem of understanding the changes that had occurred over time in the way the provisioning of goods and services was organized and drew out the implications of these changes for the structure of the American economy and for the place of the United States in the world.

In sum, with his ideas, Chandler went beyond Ronald Coase's (1988) transaction costs approach, Oliver Williamson's (1975) contracts perspective of the firm, and Stigler's (1958) original formulation about scale economies, for he found out some of the most important causes of the growth of the firm. According to Buendía (2006), Chandler's historical account needed two major improvements: it requires a formal formulation of its findings and a theory of the possibility of the firm to grow through the expansion of its economic activities to different geographical regions. So he developed he notion of *expansion economies*.

3. Determinants of the Expansion of the Firm

The theory of economies of expansion is about the relationship between the breeding of the optimal combination of all productive services of the vertically integrated firm and its output, revenues and profits. In the model, a vertically integrated firm carries out three main economic activities: purchasing of inputs, manufacturing and distribution of final products. Consequently, the firm's total profits, π, can be defined by

$$\pi = \sum_{i=1}^{N} p_i q_i - c_i q_i = \sum_{i=1}^{N} q_i \left(p_i - c_i \right) \quad (1).$$

where p_i is price, q_i is quantity and c_i is cost of the individual manufactured good i. An important assumption in equation (1) is that $p_i - c_i = \alpha$ —where α is a constant—, which implies that the firm is subject to constant returns. Given that this assumption can be introduced in equation (1) by assuming that production of good i involves a fixed cost and a constant marginal cost, it is also assumed that neither costs nor prices can be reduced.

We turn next to define the behavior of q_i, which depends of the following variables: I_i, P_i and D_i. I_i is the minimum investment or scale required to produce the inputs of the firm. We assume that if $I_i < a$, then it is neither profitable nor possible to manufacture the individual good *i*. This means that the scale of the firm's plant has to reach a minimum size *a*. P_i is the minimum investment and scale to process the inputs and produce the final good. Again it is necessary that $P_i \geq b$, where *b* is the minimal size to make production of the good of the firm possible and profitable.

D_i is the firm's distribution network, which depend on *t*, the transportation capacity of the firm necessary to bring its products to the final consumer, retailers or wholesaler; *w*, the minimum storing capacity needed to distribute final goods, and *n*, the number of outlets, stores or particular distributors. Then the firm's distribution capacity can be defined by

$$D_i = f(t, w, n, m) \geq d \quad (2)$$

The model assumes that each one of the *n* outlets has a minimum capacity or scale to function appropriately. This implies that, if the firm distributes its products through its own outlet chain, each outlet has to have the proper facilities, such as warehouse, counters, shelves, and etcetera. In the model it is assumed that the firm owns *t*, *w*, *n* and *m*, otherwise these variables are equal to zero.

The effectiveness of advertising expenditure can be described by

$$A_i = f\left[D_i(t, w, n, m)\right] \quad (3)$$

That is to say, the effectiveness of advertising depends on *t*, *w*, *n* and *m*, because while these variables increases, A_i increases as well. This is so because the increasing availability of the firm's products due to expanding distribution channels and the use of a larger number of marketing institutions, each dollar expended in advertising becomes more effective.

$$\frac{\partial A_i}{\partial D_i(t, w, n, m)} \geq 0 \quad (4)$$

In the model we assume a mutual causality between A_i and $D_i(t, w, n, m)$, because in as much as A_i becomes higher the firm the demand for its products grows so it needs to enlarge its distribution capacities.

We can define the firm's quantity of goods produce and distributed by

$$q_i = f(I_i, P_i, D_i, A_i) \quad (5)$$

Given equation (5) it is clear that the only way the firm can increase its revenue is by augmenting q_i, the quantity of goods produced and distributed. But given that q_i depends on I_i, P_i, D_i y A_i, the only way to increase q_i is by increasing I_i, P_i, D_i y A_i simultaneously y proportionally, according to the conditions established above. What we mean by proportional increases is that a new plant to manufacture inputs (I_{i+1}), the firm has to open *k* new plants to manufacture final goods (P_i). If the firm decides to open a new plant to manufacture final goods, it is necessary that it sets up *r* new outlets or stores. Obviously *k* and *r* are whole positive numbers. For instance, if a brewer decides to open a new brewing and bottling plant, it has also to set up around 600 distribution outlets.

At this point it is important to ask ourselves how much the firm will grow. As a matter of fact, the growth of the firm is determined by the local, national and international demand the firm will face. Obviously these demands have limits. For instance, a given location the firm may open three outlets. If an additional outlet makes decrease the sales of the other three outlets, the optimal number of outlets for the firms to exploit expansion economies is three. Therefore, if we divide $\overline{d}_l$ by x, the sale capacity of each outlet, and *y*, the production capacity of the firm's each plant, we obtain the optimal number of outlets at local level and the optimal number of plant the firm need at local level, respectively. As $\overline{d}_l$, $\overline{d}_n$ and $\overline{d}_i$ represents the levels of demand at local, national and international, they also provide the optimal number of plants and outlets at national and international level, given that we know the *x* and *y*.

Therefore the firm's level of manufacturing capacity expansion and level of distribution capacity are given by:

$$\frac{d_l}{x_l} = E_l, \quad \frac{d_n}{x_n} = E_n \text{ and } \frac{d_i}{x_i} = E_i \quad (6)$$

and

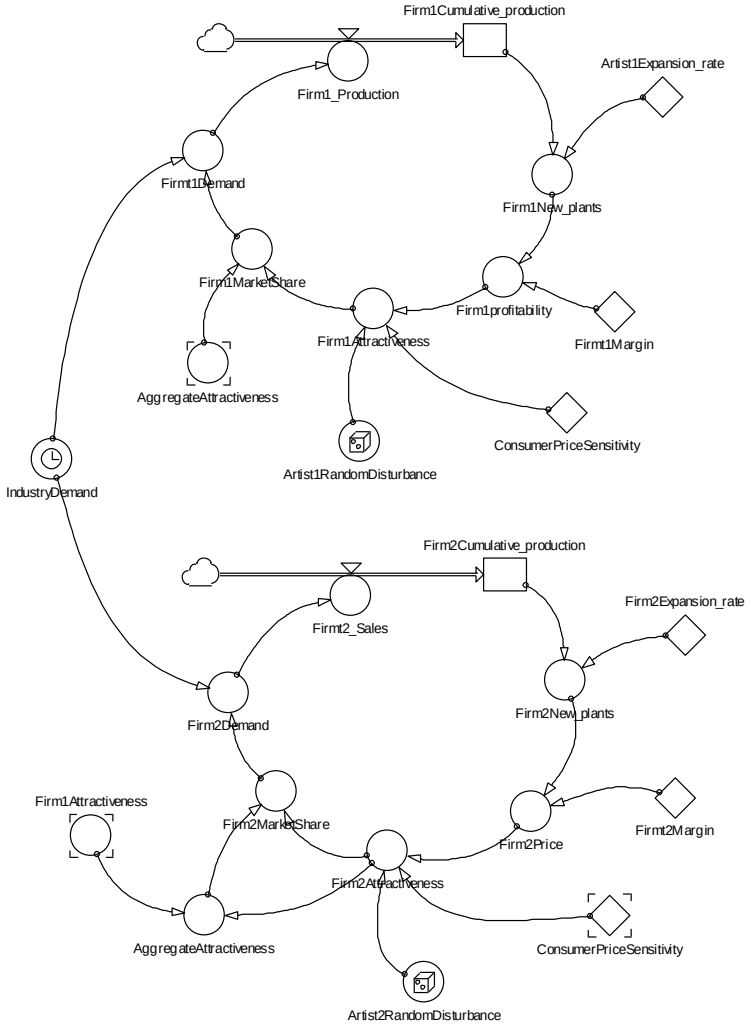
$$\frac{d_l}{y_l} = E_l, \quad \frac{d_n}{y_n} = E_n \text{ and } \frac{d_i}{y_i} = E_i \quad (7),$$

respectively.

4. Expansion Economies, Globalization of the Firm and, System Dynamics

The theory of expansion economies can be expounded with a simple dynamic model that exhibit path-dependence and self-organizing behavior. Figure 1 presents a model consisting of two firms producing an identical good. At time zero firms have the same market share. Additional market share is won through expansion economies—that is to say, by opening new plants to produce inputs and final goods and sting up outlets to distribute these goods. The primary feedback loop that affects the performance of each firm links the number of plants to produce inputs, the number of plants to produce final goods and the number of outlets with higher profits.

Figure 1. System dynamics, Expansion Economies and the Globalization of the Firm



This model has an important feature: its behavior is path dependent but not "tipping". In other words, the geographical-expansion path each firm takes is foreseen: they will share the market in a similar proportion (each firm will get 50% of the market), so a dominant firm will not emerge. This is one of the special cases where there are increasing returns and path-dependence, but where no firm will oversell the other, unless one of them has a first-mover advantage. This theory, therefore, is about that special case of markets subject to increasing returns (expansion economies) to the firm-level. Increasing returns, however, do not affect the adoption of products; therefore products in the market tend to share the market according to the geographical area they cover.

In this model, self-organization of firms is due to a network of relations that result from the mutual causality between numerous variables. As with other dynamic systems, the growth of the firm in this model is subject to both negative and positive feedbacks. Negative feedbacks produce *decreasing returns to the growth of the firm* —reductions of benefits due to scale diseconomies—, which may occur because the firm becomes "bureaucratically" congested or administratively limited. Decreasing returns to the growth of the firm are stabilization forces that hinder the growth of the firm and prevent the

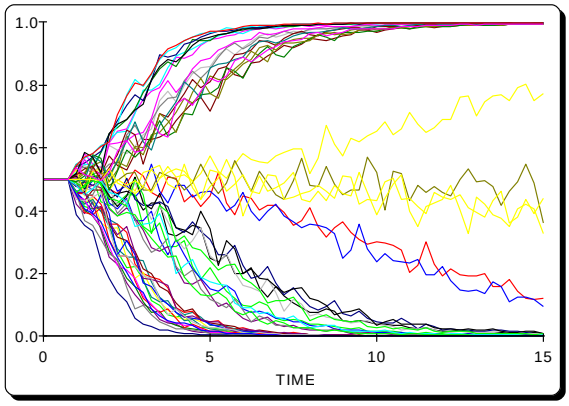
eventual emergence of an infinite-size firm. The growth of the firm and the concentration of the industry where it competes depend to a great extent on expansion economies; that is, on positive feedbacks or to put it in more economic terms on *increasing returns to the growth of the firm*.

Figure 1 presents a single sector of a simple system dynamics model that exhibits path-dependent, self-organizing behavior. The model consists of two manufacturing firms producing identical products that are competing for market share within their industry. At the beginning of the competition process, the industry demand is divided equally between the two firms. Additional market share is gained only by the opening of new outlets or manufacturing plants —in other words, the firm with more outlets or plants in any period wins more market share. Each firm can increase its production by opening new plants.

Figure 1 shows two primary feedback loops that affect the growth of firm 1. The first is a positive loop that links firm 1's stock of cumulative production to its new plants and then to its profitability, product attractiveness, market share, firm-level demand, and flow of production. The positive loop is a self-reinforcing process of new plants —that is to say, the more firm 1 expands itself, the more it sells and thus the higher its profits become and the higher the demand for its product become. The negative loop balances this growth process by the fact that firm 1's becomes increasingly bureaucratic.

Figure 1 also depicts the evolution of the market shares for each firm during four simulations of the model. Initially, the relative market shares oscillate, but at the twenty-fifth time period a dominant firm emerges. From the simulations, it is clear that the behavior of each firm's market share is self-organizing. In other words, the path each firm will take during any simulation run is not knowable from inspection of their microstructures, and the dominant firm can be different from run to run (see Figure 2).

Figure 2. Market Share Behavior



5. Conclusions

Expansion economies are an important source of increasing returns to the growth of the firm that have been neglected by conventional economics. This notion sheds light on the way the firm grows, which can be useful to revise other economic theories such as those related to international trade and industrial clusters. In this paper, using the tools of systems dynamics, I develop a model of the globalization of the firm. When a market is subject to increasing returns to the growth of the firm due to expansion economies, its structure

becomes self-organizing. This is an important result that can be helpful to revise conventional economic theories.

References

Buendía, Fernando, 2006, "Expansion Economies", proceedings of the 24th International Conference of the System Dynamic Society, July 23-27, Nijmegen, Holland

Chandler, Alfred D., 1966, *Strategy and Structure*, New York: Doubleday & Co., Anchor Books Edition.

Chandler, Alfred D., 1977, *The Visible Hand*, Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Chandler, Alfred D., 1990, *Scale and Scope: Dynamics of Industrial Capitalism*, Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Coase, Ronald, 1988, "The Nature of the Firm", Journal of Law, Economics and Organization, Spring, 4(1), pp. 33-47.

Polenske, Karen R. 2003, "Measuring Dispersal Economies", in Iréne Johansson, ed., *Entrepreneurship, Spatial Industrial Clusters and Inter-firm Networks*, Reaserch Report 01:01, University of Trollhättan/Uddevalla, Sweden, pp. 615-633.

Stigler, George J., 1958, "The Economies of Scale", *Journal of Law and Economics*, Vol. 1 October, pp. 54-71.

Williamson, Oliver E., 1975, Markets and Hierarchies, New York: Free Press.