UNEVEN DEVELOPMENT

Donald J. Harris
Professor of Economics, Emeritus
Stanford University

November 17, 2006


Author Contact: e-mail: dharris@stanford.edu; telephone: 650-725-3266.
UNEVEN DEVELOPMENT

Donald J. Harris

Abstract
Historically, one of the most striking characteristics of the general process of capitalist development is the phenomenon of uneven development, defined as persistent differences in levels and rates of economic development between different sectors of the economy. However, much of existing economic theory predicts that many of the observed features of differentiation would tend to wash out as a result of competitive market forces. This paper seeks to bridge this gap. It proposes a strategy for the analysis of uneven development that advances toward a historically and empirically relevant theory. The starting point is a conception of the firm as an expansionary unit of capital with a complex organization, existing within an ordered system of firms that constitute the global economy. The process of industry evolution that takes place in this setting manifests features of persistent differentiation arising from internal changes in the firm, interactions among firms and with the social environment, as the system as a whole undergoes development driven by the firms’ competitive activities of investment and technological innovation. The analysis draws in part on elements of the emerging paradigm of neo-Schumpeterian evolutionary theory and on some documented empirical regularities.

JEL Classification: B25, B41, D2, D4, E12, F12, L1, L2, L6, N10, N90, O1, O3, O4, O5, R11, R12.
Key words: economic growth; uneven development; competition; firms; firm competences/capabilities; entry/exit conditions; investment; innovation; technical change; industry life-cycle; evolutionary dynamics; neo-Schumpeterian theory.
UNEVEN DEVELOPMENT

In examining the general character of the process of capitalist development as it has appeared historically across many different countries over a long period of time, one of its most striking characteristics is the phenomenon of uneven development. Specifically, the process is marked by persistent differences in levels and rates of economic development between different sectors of the economy.

This differentiation appears at many levels and in terms of a multiplicity of quantitative and qualitative indices (Kuznets, 1966; Maddison, 1982; Mueller, 1990; Pritchett, 1997; Salter, 1966). Relevant measures which sharply identify the phenomenon include the level of labour productivity in different sectors, the level of wages, occupational and skill composition of the labor force, the degree of mechanization and vintage of production techniques, rates of profit, rates of growth, and the size structure of firms. This phenomenon appears regardless of the level of aggregation or disaggregation of the economy, except for the extreme case of complete aggregation – in which case, structural properties of the economy are made to disappear. For example, it appears at the level of comparing the broad aggregates of manufacturing industry and agriculture, at the level of individual industries within the manufacturing sector, and at the level of individual firms in an industry. It appears on a regional level within national economies as well as on a global scale between different national economies. In this latter context, one form that it takes is the continued differentiation between underdeveloped and advanced economies, usually identified as the problem of underdevelopment.

These disparities appear from observing the economy as a whole at any given moment and over long periods of time. While the relative position of particular sectors may change from one period to another, nevertheless there is always a definite pattern of such differentiation. We may say, therefore, and certainly it is an implication of these observations, that these disparities are continually reproduced by the process of development. Uneven development, in this sense, is an intrinsic or inherent property of the economic process. Far from being merely transitory, it appears to be a pervasive and permanent condition.

Now, it is an equally striking fact that, when we examine the theoretical literature on economic growth, we find the completely opposite picture. In particular, the dominant conception of the growth process that has motivated the post-World-War II literature is constructed in terms of uniform rates of expansion in output, productivity and employment in all sectors of the economy. In this sense, it is largely a literature of steady-state growth, whether presented in multi-sectoral or aggregative models (Burmeister and Dobell, 1970; Harris, 1978). Some notable and relevant exceptions, including Haavelmo (1954), Leon (1967), Nelson and Winter (1982), Pasinetti (1981), Salter (1965), explicitly examine aspects of the problem of persistent differentiation posed here. The recent flurry of work in endogenous growth theory seeks to incorporate some relevant elements of the problem into the neoclassical conception of the growth process (Aghion and Howitt, 1998). However, much of existing economic theory predicts that, given enough time, many of the features of differentiation which we observe empirically would tend to wash out as a result of the operation of competitive market forces (Harris, 1988). Such differentiation should therefore be viewed only as a transitory feature of the economic process.
Thus, on the one side, we find a historical picture of uneven development as a persistent phenomenon, and on the other, a theory which essentially negates and denies this fact. It is possible to go some of the way towards bridging this gap. Accordingly, I consider here a strategy for analysis of uneven development that breaks through the narrow limits of the existing steady-state theory and advances towards a historically and empirically relevant theory.

The Analytics of Uneven Development

It is necessary to start by recognizing the intrinsic character of the individual firm as an expansionary unit of capital with a complex organization. Various efforts have been made to develop a theory of the firm on this basis. (See, for instance, Penrose, 1959; Baumol, 1967; Marris, 1967, and Winter 1968/2006.) In this conception, growth is the strategic objective on the part of the firm. This urge to expand is not a matter of choice. Rather, it is a necessity enforced upon the firm by its market position and by its existence within a world of firms where each must grow in order to survive. It is reinforced also by sociological factors. It is this character of the firm which constitutes the driving force behind the process of expansion of the economy.

In the aggregate, the global economy is conceived to consist of an ordered system of firms (an interlocking network of individual circuits of capital) and its sectors (classified variously as industries, regions, national economies) likewise to be clusters of the firms that are the component units of this system. In this system, it is firms which compete, not industries, not regions, not national economies, not ‘north’ versus ‘south’. The state sets the rules and jointly determines the external conditions (externalities) within which the firms operate.

This is a crucial starting point because it establishes the idea of growth as the outcome of a process which is driven by active agents, not by exogenous factors. In particular, in the context of the capitalist economy, growth is the outcome of the self-directed and self-organizing activity of firms, each seeking to expand and to improve its competitive position in relation to the rest. Once this principle is recognized it becomes possible to move towards an understanding of the problem of uneven development.

The imperative of growth impels the firm constantly to seek out new investment opportunities wherever they are to be found. Such opportunities may lie within a wide range: in existing product lines, in new products and processes, in new geographical spaces and natural resource frontiers, or in the takeover of existing firms. However, at the core of this movement, viewed historically over the long term, are the invention, innovation and diffusion of new technologies that give rise to new products and services (Freeman, 1982; Landes, 1969, 1999; Marx, 1906, Ch. XV; Mokyr, 1990, 2002).

The emergence of growth centres or leading sectors is a reflection of this underlying process. It is a consequence of the effort on the part of many firms to create or to rush into those spheres in which a margin of profitability exists that allows them to capture new profit and growth opportunities. It may be conceived to take the form of a ‘swarm’ (Schumpeter, 1934, p. 223) or ‘contagion’ (Baumol, 1967, p. 101), marked by both entry and exit of firms. Such spheres are opened up, typically through complementary ‘macroinventions’ and ‘microinventions’ (Mokyr, 1990, p. 13) and in a sporadic and discontinuous pattern, as a consequence of the ongoing investment and innovative activity of firms and the competitive interactions among them. It is this
constant flux, consisting of the emergence of new growth centres, their rapid expansion relative to existing sectors, and the relative decline of others, which shows up in the economy as a whole as uneven development.

**The Process of Industry Evolution**

The form of this process, as it appears at the level of particular industries and products, has been identified in terms of certain empirical regularities, though there are also significant variations across industries and products. These studies show that, with some exceptions, the growth of many new industries and products follows a life-cycle pattern (Gold, 1996; Gort and Klepper, 1982; Klepper, 1997; Klepper and Graddy, 1990; Mullar-Sebastian, 1983; Wells, 1972). It may be represented schematically by an S-shaped curve of the time-path of output as in Figure 1. (For simplicity, no distinction is made here between products and processes, an industry is assumed to produce a single product, and short-term turbulence in the path of output is ignored.) Accordingly, we
may distinguish three phases of expansion: I, the initial phase, where total output is a minute share of aggregate output and grows at a low rate; II, a phase of rapid growth in which output expands rapidly and its share of aggregate output grows; III, the sector reaches a threshold beyond which its growth rate tends to level off and perhaps to decline.

To characterize the associated pattern of technological innovation, Kuznets (1979) identifies a sequence of four distinct phases constituting the life-cycle of ‘major’ innovations: (1) the pre-conception phase in which necessary scientific and technological preconditions are laid; (2) the phase of initial application involving the first successful commercial application of the innovation; (3) the diffusion phase marked by spread in adoption and use of the innovation along with continued improvements in quality and cost; (4) the phase of slowdown and obsolescence in which further potential of the innovation is more or less exhausted and some contraction may occur. This taxonomy is not all-embracing, and there are others that emphasize other features, but it is suggestive in pointing to a certain internal logic of the innovation process.

The process of industry evolution is also typically associated with a changing firm-structure of the industry. In many industries, there is a proliferation of small firms in phase I. As the diffusion of the product occurs and growth speeds up, there is a ‘shaking out’ process by which many of the smaller firms disappear (exit) and the available market is concentrated in the remaining firms. When the industry reaches ‘maturity’, in phase III, there is likely to be a high degree of concentration. This association between industry life cycle and changing firm-structure (commonly called “co-evolution”) suggests that the dynamic of expansion through innovation is simultaneously a process of the concentration of capital.

This sequence of a single product-cycle, schematically described here, is but a small segment of the time sequence characterizing the historical evolution of the economy. Given that firms are growing, making profits, and seeking to continue to grow, it must be supposed that at least some of them, having entered into phase III, would seek to launch into new investment opportunities. They will therefore actively seek to find new products that will initiate a corresponding new sequence. Alternatively, the new sequence could come from entry of new start-up firms.

It follows that we can map out the dynamic evolution of the economy as a sequential process that is discontinuous, punctuated and stochastic, with varying and overlapping time-scales of the different product-cycles, where the overall growth is accountable for on the basis of (1) the individual growth of particular new products coming on stream, (2) the growth of pre-existing products, each of which is growing at a different rate depending on the particular phase reached in its life-cycle, and (3) over time the irregular accretion of new products as the innovation process continues.

In this context, the relative position of any firm-cluster (region or national economy) at any time on a relevant index of development may be seen as a matter of the particular products it has managed to capture as a result of the previous pattern of accumulation, the ongoing activity of firms operating within it and the particular timing of their entry into the life-cycle of new products.

The causes which produce and sustain the observed patterns of differentiation must then be found within the internal dynamics of this process, leaving aside such
historically contingent factors as wars, colonial control, ‘foreign’ intervention, that may also be considered relevant and important.

What role is to be assigned to demand as a factor in this process? At the level of individual consumer products or industries, a common conception is that demand acts as an autonomous factor with a definite influence on the life-cycle pattern of evolution of the product. That influence is exerted in the early phase of introduction of a new product because of an element of resistance due to ‘habit’ formed in a customary pattern of consumption. It is exerted also in the maturity phase because of the operation of ‘saturation effects’ in consumption. But there are reasons to doubt the strength and effectiveness of such factors, as well as their supposed autonomy.

First, in an economy undergoing regular and rapid change, it is not evident what role there is for habit except for the habit of change itself. The experience of, and adaptation to, change may create a high degree of receptivity to change. What then becomes decisive in the evolution of demand (for consumer goods) is the growth of income, and the changing relative price and quality of products. Income and price elasticities of demand are an imperfect, proximate expression of this dynamic effect.

Second, insofar as these latter factors are crucial to the formation of demand, it may be argued that there is a certain self-fulfilling aspect of the expansionary process at the level of industry demand. In particular, investment generates the demand that provides the market for the new products which investment itself creates. This occurs in two ways. One, structural interdependence in the economy at the level of both production and expenditure patterns, allows for the possibility of a certain mutual provisioning of markets when expansion takes place on a broad front. Two, as a new product unfolds through the stages of the innovation process, it undergoes both improvements in quality and a decline in price relative to other products. This development provides a substantive basis for making inroads into the market for existing closely related products and hence promotes demand through a shift from ‘old’ to ‘new’ products. It is perhaps this shift effect which is mistakenly identified as a saturation effect by adopting a one-sided and static view of a dynamic and interdependent process.

Each and every individual firm must of course secure a place in the market for its product. Its success in this regard is dependent on its own efforts and capabilities.

**Competition, Firm Capabilities, Entry/Exit Conditions, and the Social Environment**

Analytical treatment (including formal modeling) of the process of industry evolution has flourished in recent years in tandem with an outpouring of empirical studies covering different industries, countries, and time periods. Much of this work is done within the frame of an emerging paradigm in the Schumpeterian tradition of evolutionary dynamics (Futia, 1980; Iwai, 1984a, 1984b; Nelson & Winter, 1982; Dosi, 1984) and there are other theoretical approaches (Lowry, 1979; Dasgupta and Stiglitz, 1980; Durlauf, 1993). For a review of the current state-of-the-art and challenges for research, focusing on the evolutionary approach, see Malerba (2006). Relevant for present purposes are the significant insights provided so far by this work into the mechanisms and causal factors which govern the process of industry evolution and account for the persistence of differentiation among firms.
The Neo-Schumpeterian approach develops an explicit formulation of ‘Schumpeterian competition’ in which firms innovate to win super-normal profits, profits are reinvested to provide further growth through innovation and market expansion, and there are winners and losers due to the operation of selection mechanisms and learning mechanisms. Decisions are typically based on bounded rationality. It is shown that such competitive behavior under specified conditions gives rise to persistent differentiation among firms in terms of size, productivity, costs of production, product characteristics, profitability and growth and may breed long term sustainable market concentration among surviving firms, with or without entry. Economies of scale and scope are not a necessary part of the story; a key factor is increasing returns to knowledge and learning. Though there exists a strong tendency to concentration, it is not inevitable, and depends on industry characteristics that vary across industries. There also exist dual tendencies of ‘creative destruction’ and ‘creative accumulation’.

A distinctive feature of this approach is the conception of the firm itself as an organizational unit. The firm is conceived as the embodiment of a set of strategic assets (competences or capabilities), tangible and intangible, consisting of knowledge, skills, and routines, gained through path-dependent experience and learning, that are specific to each firm and non-tradeable. These assets evolve over time (through ‘competence accumulation’) with the ongoing process of evolution of the industry and through interaction with the changing environment. Consequently, diversity among firms is not only a characteristic of the system of firms, it is also reproduced by the evolutionary dynamics of the competitive process.

Some key factors determining the evolutionary path of industry structure in terms of firm composition are the following. (1) First (second, third)-mover advantages arising from a combination of unique internal attributes of the mover, product characteristics, network effects among users, and random chance events. (2) Non-pecuniary network externalities associated with cues and information gained from interacting with the ‘local’ social environment of firms, users of the product, and institutions involved in knowledge creation and information dissemination (on the national level, the ‘national system of innovation’). (3) Spillover effects among firms and across industries, which may be both positive and negative. (4) Increasing returns to knowledge and learning within the firm. (5) A firm may become ‘locked-in’ to its own trajectory of technology development and reap increasing returns therefrom, but eventually suffer a disadvantage from generating irreversibilities and inertia causing inability to adjust to change (‘success breeds failure’). (6) The very same factors which confer advantages upon early entrants and incumbent firms may create barriers to entry for ‘latecomers’, depending on the stage of industry evolution and timing of entry.

Some relatively neglected factors that need to be integrated into a more comprehensive analysis include: (1) the role of market demand, as related to the mutual interaction between producers and users (consumers, other firms, and the state); (2) the role of the financial system (Schumpeter had assigned a crucial role to the granting of credit ‘as an order on the economic system to accommodate itself to the purposes of the entrepreneur’ (1934, p. 107)); (3) workplace and labor market interactions, lightly touched upon by Mansfield (1968, ch. V) and vividly described in historical detail by Braverman (1974); (4) the system of governance by the state, that sets and enforces the rules and norms, including property rights, governing conduct by firms.
Within this extended framework of analysis, it is possible to explain not only how some firms (or firm-clusters) come to capture the position of leaders (and may eventually lose it to others), but equally how some are left behind, others drop out altogether (exit), and still others remain on the ‘periphery’ (so to speak) lacking the internal and external capabilities to enter. In this regard, the explanatory power of this analysis is readily applicable to commonly discussed empirical and historical phenomena such as ‘deindustrialization’, ‘catching up’ (convergence), and ‘falling behind’ (divergence).

What emerges from this analysis also is an understanding of the critical role of public policy and programs to foster economic development. Because of the pervasiveness of externalities and various forms of coordination problems, market failures are intrinsic to the process, calling thereby for collective intervention to achieve efficiency and socially optimal results.

**The Aggregation Problem**

All of the preceding analysis concerns the pattern of industrial growth viewed at the level of an individual industry and the firms (or firm-clusters interpreted as, say, regions or countries) that compose it. There is nothing in this analysis to indicate how the pattern of growth of different industries translates into aggregate expansion at the level of the economy as a whole, or how the various industrial patterns fit together to form a complete whole. This is a substantive problem requiring further analytical treatment on its own terms. Its significance derives from recognition that the economy as a whole is not just the sum of its parts. Hence, the motion of the economy cannot simply be deduced from the movement of its parts. The usual methodological device of the ‘representative firm’ necessarily fails in the present context.

A related aspect of the problem is associated with the manifold and complex ways in which growth in one sector (however defined) mutually conditions and is conditioned by growth in other sectors. Such mutual interaction is a necessary consequence of economic interdependence in both production and exchange. (Hence, models of international trade that claim to show uneven development arising uniquely from exchange of products give a one-sided representation of the problem.) The existence of such interaction implies that there is a certain cumulative effect intrinsic in the growth process. Understanding the exact mechanisms through which this effect operates is one of the central analytical problems for the analysis of uneven development.

There is no guarantee that in the aggregate there is always sufficient demand for all products. It is here that the analysis comes full circle, back to the problem of overall effective demand that motivated the early post-war growth theory initiated by Harrod (1948) and Domar (1957). This problem was a central focus of the analysis in the Keynesian and Post-Keynesian tradition, less so in the case of the neoclassical tradition (as detailed in Harris, 1985). It appears now that it cannot be escaped in making the transition to the analysis of uneven development.

The analytical framework presented here lays the groundwork for addressing this larger set of problems.
Bibliography


