Why Do Firms Differ and How Does It Matter? A Revisitation

Richard R. Nelson *

It has been seventeen years since I published a paper (Nelson 1991) concerned with the determinants and consequences of firm differences. I have been pondering the question of whether I now have anything useful to say about it that I did not say then. I don't know, but let me give it a try. I begin by reflecting on what we might learn from studies of biological evolution about the importance of narrative and qualitative observation in analysis of the factors behind and consequences of firm differences. Then, I turn to commentary and criticism of what I think has been the dominant point of view (certain authors clearly are exceptions) orienting the studies of firm differences and industrial dynamics using the available longitudinal industry-firm data sets. Finally, I lay out some of my own rethinking regarding the sources and consequences of firm differences.

Keywords: Firms, Variation

JEL Classification: D01, D21

I. Studies of Biological Evolution and the Importance of Narrative

To begin, I want to call attention to some aspects of biological evolution. While I am on record as insisting that economic evolution differs from biological evolution in several essential ways, in this case at least I think we can learn something from biology. So let us reflect on why there is such variety among entities that are classed as animals,

* Professor Emeritus, George Blumenthal of International and Public Affairs, Columbia University, New York, NY USA, Visiting Professor University of Manchester, Manchester UK, (E-mail) rrn2@columbia.edu. This paper was originally presented in a conference on firm differences, hosted by Maison Franco Japonaise, held in Tokyo in November 2008.
[**Seoul Journal of Economics** 2008, Vol. 21, No. 4]

or more narrowly as birds, or still more narrowly as finches, or among finches on a particular island in the Galapagos.

The first thing to note is that except perhaps for the last, the categories are quite broad. While there are certain characteristics that all entities in the class must share — indeed some of these characteristics define the category — one should not be surprised by considerable within category variety. And this certainly also is so for firms included within an industry, as defined by the statistical agency that collects and publishes the data. The expected heterogeneity of course is greatest at the most coarsely defined industry level, but the industries in even the finest cut clearly contain a variety of different kinds of firms, in different circumstances, catering to different groups of customers, doing different things.

Turning to finches, there are a variety of environments within which finches live, and even some variety within particular islands in the Galapagos. Some of the birds are chicks, some in the prime of life, a few quite old and approaching the end of their life spans. Some of the mature birds were undernourished when young, and some had accidents that left them handicapped in some way. These differences are likely to be correlated, but not necessarily strongly, with ability to live through another year, or to mate successfully.

Similarly for firms, even for those contained within an industry defined at a very fine level, as say grocery stores in a particular suburban area of a particular city. While there likely is overlap of markets, different firms are located in different places, and are convenient to different local neighborhoods. Some have a reputation as having good butchers, others don't. Some customers are particularly interested in buying good meat; others are vegetarians. Some stores have been around for a long time; some have just started. Among the former, some have had the same ownership and management for a long time; others are under new management, or have recently been acquired by new owners, or both. This is so both for grocery stores that are members of a national chain, and those that are owner operated.

The grocery stores also certainly differ in the profits they earn, and in their total factor productivity. There almost surely would be some correlation between these variables and the firm characteristics mentioned in the paragraph above. However those correlations are not likely to be high, even if there were good measures of profit and TFP for each firm, which there surely are not.

And the correlation between these variables and survival and growth of a firm is unlikely to be very close. Some stores operating with low productivity and negative reported profits have been just started, and have the backing of relatively patient money. Other stores with low accounting profits are owned and operated by a family who expect from the store only a modest living. The owners of some highly profitable stores may be quite uninterested in starting a new branch, or in expanding their operation and responsibilities in other ways. To make a good guess of whether a firm survives to the next period or not, and if so its growth rate (plus or minus), almost surely would require that the observer-predictor look at more than the easily available numbers describing each firm.

Biological evolution is marked by changes over time in the phenotypes that are dominant in a species, with the changes often roughly describable in terms of a series of punctuated equilibria. The change in the dominant phenotypes from one equilibrium to the following one often can be ascribed largely to changes in the environment that supports and challenges population members, without any significant role played by the introduction to the population of new types. However on other occasions the key force upsetting the old equilibrium and initiating change can be attributed to a burst of mutation which created new phenotypes even better suited to the existing environment than the previously dominant forms.

Similarly in industrial competition. A change in the needs or preferences of customers or of input supply conditions may have a major effect on the viability of firms in an industry that are oriented differently regarding product line or production process. While the distribution of firms may change significantly the character of individual surviving firms may not. On the other hand, as Schumpeter argued long ago, change in industry structure in many cases is driven by the introduction by one or a few firms of a new product or a new production process which suits the existing market environment better than what had been prevailing dominant practice, and the innovators and their close followers may emerge as industry leaders.

Is firm growth associated with successful innovation? In the second case yes, but in the first case no.

This summer I reread Marshall's Principles, and for the first time read his Industry and Trade. In the parts of Principles that deal with firms, Marshall does make extensive use of the concept of a "representative firm." However, he is very clear in his articulation that he uses

the concept in order to discuss in a compact way things that are going on in the industry as a whole, and that the reader needs to recognize that firms in an industry are very diverse. And particularly in his Industry and Trade, Marshall characterizes the environment within with firms operate as dynamic not static, and argues that the firms themselves, at least the ones that are successful over time, generally are changing. Stanley Metcalfe recently has been arguing that Marshall in fact was an evolutionary economic theorist.

I think it highly relevant for the discussion we are having here that Marshall as an economic theorist — evolutionary, neoclassical, or both — clearly believed that the economic reality was too complex for mathematics to be the appropriate language for economic theory, at least of the level of abstraction that Sidney Winter and I have called appreciative theory. In his Principles Marshall relegated his use of mathematics to footnotes and appendixes, and his use of mathematics there was largely to sharpen up some of the propositions he had made verbally in the text. The overall theoretical picture is presented verbally. And his description of industry in his Industry and Trade is largely a verbal picture. Numbers — statistics — are sprinkled throughout the verbal description to give it hard form in places.

It is interesting, and I think relevant to this discussion, that evolutionary theory in biology is largely articulated verbally, and the empirical phenomena that evolutionary theory addresses largely described in narratives accompanied by pictures and charts. This is so for advanced treatises as well as in more introductory texts. Of course there are a number of mathematical models used in contemporary expositions of evolutionary biology, John Maynard Smith's game theoretic formulations perhaps being the most well known. I once had the occasion to talk with Maynard Smith about what he thought his models were about. He was emphatic that they did not amount to a way of formulating evolutionary theory; as he put it "evolution is much more complicated than that." Rather, he argued, that they were useful as aids and stimuli to thinking, a complement to the broader less formal articulation of evolutionary theory.

I would like to propose that Maynard Smith's point of view here on evolutionary biology strikes me as very similar to Marshall's on economics.

Why am I making these points? Mainly to argue that the research by economists on differences among firms and the consequences of these differences for firm survival and growth has tended to overemphasize

formal modeling, and quantitative analysis of a few firm features, and that better understanding of what is going on here is likely to require a richer and more complex theoretical formulation, and more detailed observations of firms including a variety of their qualitative features. My argument here most emphatically is not that formal dynamic modeling isn't a useful tool for analysis of the questions in which we are interested. Nor is it to denigrate efforts to develop a variety of measures of firm characteristics and performance to use in econometric exploration of these questions. Rather, it is that these tools of analysis need to be imbedded in, and used to complement, detailed case studies of firms and narratives of what has been going on in an industry. And that by and large this has not been done.

II. Studies of Firm Differences and Industrial Dynamics Using the New Data Sets

Shortly after Sidney Winter and I finished our An Evolutionary Theory of Economic Change (Nelson and Winter 1982) I began to come aware that data sets of firm and industry dynamics were beginning to become available. These data sets — many at this gathering have worked with them — provided not only a cross section at various periods of time of firms or establishments within an industry, but also a tracking over time of the individual units. Or it looked as if these data sets would enable such an analysis.

I salivated at the prospect. There seemed to be the material here for a convincing documentation of many of the basic premises of evolutionary economic theory, and the opportunity to explore and find out about key variables and relationships in that theory. For a variety of reasons (partly my own lazyness), I never myself have dug deeply into these data sets. But I have followed with great interest the now quite numerous studies of this genre.

As I think we might all agree, perhaps the most striking results of these studies are the following. First, there is totally convincing evidence supporting the proposition that at any time there is great variety among the firms in a nominal industry, in terms of measures of how they are producing things, productivity, and where we can construct them measures of profit, and also considerable variance of firm or establishment growth rates. But second, the simple evolutionary model that proposes that profitable firms expand and unprofitable ones contract

does not fit the data very well; the dynamic processes involved clearly are much more complicated than that model suggests.

Or at least the dynamic processes in the data sets being analyzed. In my earlier optimism regarding what could be explored in these kinds of study, I clearly had only a weak understanding of the broadness of the standard industrial categories, even the fine grained ones. I believe I and many others, on our first thinking about what we might expect regarding the factors influencing the survival and growth of firms did not appreciate adequately the range of different markets faced by different firms, and other differences that set firms apart, and in appropriately had in mind a set of firms in active competition with each other, along the lines of those simple models.

It is important to recognize that the problem here is only partially characterizable in terms of the impurity of the industry classifications. Particularly in parts of the economy where there is significant product differentiation, it is conceptually impossible to define sharp disjoined industry categories based on markets. For some customers cars and motor-cycles and pick-up trucks are alternatives they consider when buying a family vehicle, but for many customers these are quite different commodities. For some a Chevrolet and a Mercedes are competitive candidates for purchase; for others they are very different kinds of cars. Schumpeter's characterization of competition in industries where product innovation is important implicitly denied the sharpness of industry boundaries and recognized the fact that customers often differ in what they want of a product and hence what competes with what.

In doing our theoretical, as well as our empirical, analysis, evolutionary economists need to recognize better than we have that heterogeneity of markets within an industry is the rule not the exception, and that the lines between one nominally defined industry and another may be blurry not sharp.

I think there also has been a tendency to oversimplify the relationship between firm or establishment profitability or relative efficiency compared with its competitors, and survival and growth. Of course as indicated above there is, first of all, the question of which firms are real competitors and which ones are not. Then there are major issues about the adequacy of the measures of firm efficiency that we can calculate. But these issues aside, as the examples I mentioned in the preceding section signal, there are good reasons to doubt a strong relationship between efficiency or profitability and growth, at least in the short run. Some firms may be just getting their legs under them-

selves and have the resources to keep going, even growing, while not nominally profitable. Other firms may have high profits selling in a particular market, where demand is relatively inelastic and expansion of current activities not likely to enhance profit.

I want to argue that special attention ought to be given to identifying and analyzing new entrants to an industry. My strong suspicion, based on some of the arguments above and the several studies that have been made of the fate of new entrants (in general not based on the "complete" industry firm sets under consideration here) is that, after normalizing for firm size, entrants have a greater variance of most firm characteristics, a higher tendency to fail, and a greater variance of growth rates, than established firms.

It is my belief that the variety and turbulence in an industry at any time is related to the relative importance of new entrants and their characteristics (for example the size at which they enter). However, to my knowledge few if any of the studies in question have looked at this matter.

Many, not all, of the economists studying firm differences and industry dynamics are Schumpeterian, in the sense that they see continuing innovation as the key driving force behind economic change and the reason for continuing disequilibrium. I think it fair to say that analysts of this ilk tend to have in their heads a model in which in general firm innovation tends to be profitable, a successful innovation tends to show up in higher than average total factor productivity (TFP) (or an indicator of that) of the innovating firm, and successfully innovating firms tend to expand their share of the activity in the industry as a whole, while firms that do not innovate and do not imitate quickly tend to experience a declining share. But empirical studies of industrial innovation, and of innovating firms, tend to highlight that the success of an innovation is highly uncertain. In many cases innovations lose money. And particularly when an innovation is first introduced, there are likely to be a variety of problems with it that may be associated with low TFP.

Many years ago Mansfield (1962) did a study in which he reported that the firms he identified as innovators in an industry tended to grow more rapidly than firms that were not innovators. I have not followed recent work that has followed along Mansfield's path. However, I confess being suspicious regarding whether there is any simple relationship between firm innovation and firm growth.

Some innovations fail, and the firms introducing them may be

seriously hurt by their efforts. Other innovations succeed, but often with a lag. While to my knowledge this has not been subject to careful empirical investigation, I would conjecture that, as an innovation begins to succeed, profitability and productivity tend to go up, and the innovating firm expands, largely because the potential demand for the innovation, either as a marketed product or as a production process, has become clearer and more manifest.

If this is a reasonable (if oversimplified) characterization of industrial innovation and of what happens to innovating firms, then one would not expect to see a close relationship among an indicator of whether a firm has innovated recently or not, and TFP, or firm growth. Rather, one would expect to see considerable variation among firms that have innovated both in the various measures of their performance and in their growth rates. Within that variety, one might expect to see a correlation between growth of TFP of a firm and its growth rate. Anyhow, I propose that this is a more plausible relationship to look for than those that have been explored in many studies, and found not to hold very well.

One of the most basic understandings that has been won through research on industrial innovation is that industries differ significantly. They differ in the rates of innovation experienced, and in the type of innovation that is most common. Technological innovation in semiconductors is, by all reasonable assessments, much more rapid and significant than in shoe production. In some industries product innovation is the driving competitive force; pharmaceuticals is a canonical example. In other industries, most of innovation involves the improvement of processes of production which, in some cases, involve improvements in product quality, but there is little product innovation per se. The steel industry is a good example, as well as much of textile production. In some industries most of the innovation comes from firms in the industry; in others firms outside the industry, often suppliers of equipment or materials inputs are the principal sources. In some industries, or for periods of time in the lives of those industries, innovations largely come room established firms; in other industries or in other periods of times, new entrants play a major role.

Thus there is good reason to believe that the critical factors behind differences among firms in an industry at any time, and the consequences of those differences are not the same in all industries. I propose that identifying and analyzing differences across industries ought to be high on the agenda of research in this area.

Over the same period of time that the new longitudinal firm-industry data sets became available, there also has come available considerable data about the characteristics of technological innovation in different industries. Partly this new information has come from a variety of "innovation surveys," which have probed at various of the dimensions mentioned above, and others, like the mechanisms through which firms in an industry try to gain returns from the innovations they make, and their effectiveness (For the pioneering study see Levin *et al.* 1987). To my knowledge, no one has tried to combine these two sources of information. The illumination could be great from doing so.

And let me return to my earlier argument about the importance of using detailed firm case studies and industry narrative as an important part of our understanding about the kind of topics we are considering here. I would argue that these provide a potentially very valuable, and neglected, body of knowledge for posing the important questions about firm and industry dynamics in an industry, and for interpreting the statistical findings.

III. Some Rethinking Regarding Firm Differences

While the now numerous studies of firm differences and firm and industry dynamics, using the now not-so-new data sets, have not reinforced the simple story about the processes of economic evolution that I once had in mind, I think I have learned a lot from them. I want to conclude this essay by highlighting some considerations bearing on the factors behind firm differences and the consequences of those differences that I now see more clearly, or at least differently, than when I wrote that original essay on this subject.

For one thing, I now have a much better appreciation than I did that a significant amount of the differences among firms in a nominal industry reflects the fact that different firms are serving different needs and different groups of customers. The nature of this kind of firm variation within a nominal industry may not show up clearly in the standard data sets. It is likely to be reflected in variation in firm sizes, as a result of differences in the sizes of the sub-markets they serve. But it is unclear whether this kind of variety is associated with differences in input coefficients or in total factor productivity. Measures we construct of firm profit rates are not very reliable. In any case, the firm variation resulting from different market niches can be rather

stable. In industries where this is the dominant cause of firm variation, and there is not rapid change in underlying technologies or the pattern of demand, I would not expect to see any strong relationship between the firm characteristics we can measure reliably and firm growth rates.

I would argue we need to recognize this source of firm variation better, and distinguish it from firm variation associated with the forces of economic change. On the other hand, I do think that much of firm variation does have to do with the dynamic processes of change, and I continue to believe in the value of understanding these processes as evolutionary.

A large part of the motivation for an economic evolutionary theory of economic change was and is to recognize explicitly that the flexibility of a firm at any time often is quite limited. However, obviously firms are not stuck with their routines in the same sense that phenotypes are stuck with there genes. They are locked in neither to their operating procedures, nor to their size. A basic question for evolutionary economists to explore is the relative importance in the advance of performance at an industry level, say measured in terms of industry productivity, of on the one hand the growth of productivity of the individual firms, and on the other hand the expansion of high performance firms relative to firms with weaker performance (including entry and exit). Before the results of the studies in question began to come in, while I strongly suspected that the relative importance of these two mechanisms differed from industry to industry, I did not quite know what to expect here. However, I confess being surprised by the weakness of the latter mechanism in virtually all the studies, and the dominant role played by the improvement in the productivity of a significant fraction of the firms in the industry.

I shouldn't have been surprised. For some time I have been arguing that an important aspect of economic evolution, another factor that has no real counterpart in biology, is the important role of knowledge that is widely shared by firms in an industry, and which tends to generate a certain similarity of prevailing practice, and often also a broadly similar orientation of efforts at invention and innovation. This common body of knowledge, which Sidney Winter and I associated with a "technological regime" and Giovanni Dosi has called a "technological paradigm," has three different, if overlapping, sources. First, awareness of firms in the industry of what their competitors are doing. Second, in most industries there are professional societies that include the technical people in the different firms, and provide a vehicle for the

sharing of information. Third, many technologies are associated with underlying fields of scientific research with largely open publications and meetings.

Of course not all technological paradigms are of the same strength; obviously their strength differs significantly from industry to industry, and within an industry can vary over time.

There is the interesting and important case of technologies that are new, and where beliefs about what is needed to make products or processes based on it technologically and economically viable are still very much in flux. Thus in the early days of the automobile, or the modern computer (to pick two canonical examples) there certainly was a body of knowledge and belief shared among those working in the field. However, the embryonic paradigm was very loose. Beliefs about what were the likely most important uses of the new technology, and of the designs that could best meet those latent demands, varied widely. Different firms made different bets on these matters. These basic questions were answered in good part through the variety of different designs that different firms and their customers tried out, and experience with and feedback from operating experience with these.

In this process, the body of know-how shared by professionals in the industry grew more sophisticated and powerful, and firms learned from the successes and failures of their competitors as well as their own successes and failures. However, as best I am able to read these and similar histories, a large part of the story of the emergence of a viable industry involved the success and growth of a few firms and the failure of many others.

It is not surprising that this picture is not the one that is showing up in most of the empirical studies of industry dynamics. I would argue that the industries in these studies are not ones that are just emerging (although certain sub-sectors of them may be new). Rather, they are ones in which broad technological paradigms are established, and play a major role in molding industrial dynamics.

In some of these industries, the paradigm may be relatively loose, leaving considerable room for differences among firms in what they think are the practices, including policies towards innovation, that are going to prove to be profitable. One would expect to see in such industries a reasonable amount of firm variety, including significant differences among firms at any time in their efficiency and profitability.

On the other hand, I suggest that even in industries where the paradigm is relatively loose, in most cases it is substantial enough to

prevent the winners in the innovation competition from completely outdistancing their rivals, and eliminating them from the market. Firms that have not successfully innovated in general will be able to learn from their rivals' successes, and themselves do something similar. In such a context then one would expect to see just the picture of many firms increasing their productivity that we have seen in the various studies. Some of these firms are imitators, or at least followers, rather than the original innovators. And because many lagging firms are able to respond often relatively quickly to the leaders with advances of their own, the fraction of overall industry productivity growth accounted for by the growth of firms of higher than average efficiency and the decline of those with less is relatively small. Or at least that is my proposal here.

Not surprisingly, I propose that in industries where the paradigm is strong, even more so than in ones where the paradigm is weaker, one would expect to find that advances in the productivity of individual firms accounts for almost all of industry level productivity growth. In such contexts, I would suspect that the growth and decline of firms would have little connection with their relative efficiency rankings. These remarks are not meant to downplay the importance of competition in such industries. However, in industries where the technological paradigm is strong, the importance of competition is largely to spur firms to continuingly advance their technologies, or lose out to their more innovative colleagues. However, I would propose that the importance of competition in generating variety in such industries is less than in industries where the paradigm is looser.

Of course technological paradigms are not static things. They, as well as industry practice, and the structure of industry, evolve over time. Sometimes the rate of advance of a paradigm is very rapid, particularly in technologies that are associated with progressive fields of science. But a characteristic of many technological paradigms is that while they progress, they tend to preserve their basic intellectual structure for long periods of time.

On the other hand, while a dominant paradigm can last for a long time, and enable continuing progress, economic history suggests that they seldom last forever. Old paradigms tend to run into diminishing returns ultimately, and new ones emerge that are seen by some as having considerable promise. When the latter happens, the forces that bind firms in the industry together tend to weaken, new firms enter or old firms commit to trying something radically new, or both. In effect

we have a new industry. As I suggested above, in contexts such as these variety per se is extremely important.

Taking account of the difference in industries I have been suggesting above regarding the nature and significance of firm variation and industry dynamics of course would require that analysts of the data sets under discussion orient their analyses, and interpret their findings, in the light of industry characteristics that one can see best from more qualitative detailed narratives regarding what has been happening in the industry. The case for bringing together of these different kinds of information and different kinds of ways of characterizing industry dynamics is perhaps the central one I want to make in this essay.

(Received 6 November 2008; Revised 16 December 2008)

References

Levin, R., Klevorick, A., Nelson, R., and Winter, S. "Appropriating the Returns to Industrial R and D." *Brookings Papers on Economic Activity* 18 (No. 3 1987): 783-820.

Mansfield, E. "Entry, Gibrat's Law, Innovation, and the Growth of Firms." *American Economic Review* 52 (No. 5 1962): 1023-57.

Marshall, Alfred. *Industry and Trade*. McMillan Press, 1919.

Nelson, R. "Why Do Firms Differ and How Does it Matter." *Strategic Management Journal* 12 (Winter 1991): 61-74.

Nelson, R., and Winter, S. *Evolutionary Theory of Economic Change*. Cambridge: Harvard University Press, 1982.