Networks, Competition, Innovation and Industrial Growth

Hans W. Gottinger

Economic Issues, Problems and Perspectives

NOVA

ECONOMIC ISSUES, PROBLEMS AND PERSPECTIVES

NETWORKS, COMPETITION, INNOVATION AND INDUSTRIAL GROWTH

ECONOMIC ISSUES, PROBLEMS AND PERSPECTIVES

Additional books in this series can be found on Nova's website under the Series tab.

Additional e-books in this series can be found on Nova's website under the e-book tab.

ECONOMIC ISSUES, PROBLEMS AND PERSPECTIVES

NETWORKS, COMPETITION, INNOVATION AND INDUSTRIAL GROWTH

HANS W. GOTTINGER

nova publishers
New York

Library of Congress Cataloging-in-Publication Data

Names: Gottinger, Hans-Werner, editor.
Title: Networks, competition, innovation and industrial growth / editor, Hans W. Gottinger.
Description: Hauppauge, N.Y. : Nova Science Publishers, Inc., 2015. | Includes index.
Identifiers: LCCN 2015039938 | ISBN 978-1-63484-022-4 (eBook)
Subjects: LCSH: Business networks. | Competition. | Industries--Technological innovations.
Classification: LCC HD69.S8 N475 2015 | DDC 658/.046--dc23 LC record available at http://lccn.loc.gov/2015039938

Published by Nova Science Publishers, Inc. † New York

CONTENTS

PREFACE

"The financial crisis raises a potentially existential crisis for macroeconomics. Practical macro is based on the assumption that there are fairly stable aggregate relations, so we do not need to keep track of each individual, firm, or financial institution – that we do not need to understand the details of the micro plumbing.

We have learned that the plumbing, especially the financial plumbing, matters: the same aggregates can hide serious macro problems. How do we do macro then?"

(Olivier Blanchard, Chief Economist IMF, 2015)

In his 'Lever of Riches' (1990), noted economic historian Joel Mokyr argued about the sources of ideas as a precursor to technological development in the Renaissance period in Europe as those making their way from China, India, Persia, and the Middle East. In the period after 1500, the gap in producing new ideas for technological progress between Europe and the other regions in the world widened and a few centuries later it led to the eurocentric 'Industrial Revolution.' Industrialization itself in Europe created leaders and followers and spread to then eurocentric North America with changing leadership all the way, invoking catchup development in technology, industry and the economy.

What it teaches us beyond history is that we could identify a chain of precursor events and stages, gradually evolving and expanding on ideas, inventions, technologies, innovations, industries, networks, markets, competition and industrial growth. All of these events will mutually reinforce the dynamics of this process; networks will be a major facilitator, and since it works both ways, they will act as a central catalyst.

We attempt to embed and synthesize the chain of events as the building blocks of industrial development.

This text should be of interest to policy analysts/makers, industry strategists/consultants, students of industrial economics/organization, entrepreneurship/management and economic journalists alike with interest in and focus on strategic and structural foundations of network economies and their practical implications. The book is also designed to be used as a text for a course in business strategy, as it could serve as supplementary to industrial organization and to the micro-foundations of economic growth and development—a theme that has been advanced by F. M. Scherer in his "Innovation and Growth" (1984). It was proposed in a sequence of follow-up studies, but neglected in the mostly macro -centered economic growth literature.

Some structural elements of networks carry over to features of dynamic competition in network industries that, through increasing return mechanisms on innovation based industries, generate sustainable growth and create industrial links as a backbone to industrial and economic growth. The project focuses on the integration of network structure, industrial competition and economic growth processes into a coherent mechanism design.

The central theme of this compilation is the interplay of competition, innovation, cooperation and market structure along vertical and horizontal industry lines. Interaction and interdependencies are facilitated and sometimes bottlenecked through networks most naturally prevalent in high technology industries. This forms the core basis of business strategy relating to the growth of business and complementary activities through innovation, mergers and acquisitions (M&As) and related strategic choices.

As Edith Penrose mentioned in her seminal monograph "The Theory of the Growth of the Firm" (1959, 1995) "...competition is at once the god and the devil," we also look at likely aberrations of competition in the marketplace (antitrust).

The selection of the material springs from research and consultancy work in ICT related network industries with STRATEC Munich, www.stratec-con.net, and further developed from previous books by the author such as "Economies of Network Industries(2003)," "Innovation, Technology and Hypercompetition "(2006), "Strategic Economics of Network Industries" (2009) and with M. Goosen (eds.), "Strategies of Economic Growth and Catch-Up" (2012).

Many of the ideas presented here emerged from conversations with friends, colleagues and former collaborators on these issues. In particular, I want to mention Profs M. Takashima (Yokohama) and C. Umali (Nagasaki). Discussions with participants at various conferences such as the J. B. Say Conference on Innovation in Boulogne-sur-Mer in 2014, the International Conferences on Emerging Industries (ICEI-13, ICEI-15) in Shenzhen, China and the European Investment Conference (EIB) in Brussels, Nov. 2014 were very helpful.

Ms. Juliane Sonntag, Vienna provided the figures in Chapters 2 and 4.

My sincere thanks go to all.

Hans W. Gottinger, June 2015

Chapter 1

INTEGRATING NETWORKS, DYNAMIC COMPETITION AND INDUSTRIAL GROWTH

We look into the integration of structural characteristics of economically relevant industries with innovation driven dynamic competition and strategies for micro- and macroeconomic growth. Some structural elements of networks carry over to features of dynamic competition in network industries, further facilitated through increasing returns mechanisms in innovation industries that generate sustainable growth by creating industrial value-based linkages for backbone industries to economic growth. Further on, in Chapter 6, this work not only focuses on the integration of network structure, dynamic industrial competition and economic growth but also on its transfer to a coherent economic mechanism design. Here the overall central theme is the interplay of competition and cooperation along vertical/horizontal industry lines (Gottinger, 2009). This lays the groundwork for business strategies relating to the growth of business and complementary activities through innovation, mergers and acquisitions (M&As) and related strategic choices (Besanko et al., 2000).

There are intrinsic links between networks, dynamic competition and economic growth. In conventional organization theory typical network industries are those where physical networks are dominant as in energy power grids, transportation, airlines, railways, roads, telecommunication and communication lines.

But virtual networks encompass every industry and the public at large in a variety of smaller or larger networks. No industry within a standard industry classification (SIC) stands alone. Almost everything is connected. This is the name of the game for the 'Internet of Things' (IoT) or what is sometimes called 'Industry 4.0.' As the economy advances in a technology and network driven direction it will become a reflection of a highly connected Internet (see Case A).

From a country's industry perspective, some industries would form major hubs (platforms), with a few hub nodes (vertices) with links (edges) to other nodes possibly partitioned into smaller local hubs. Others form smaller network clusters that extend beyond borders. Further, strongly innovation and technology driven industries will dynamically extend links to new entities that essentially form new markets. They are offsprings of technological opportunities or organizational systems that relate to new businesses as newly formed increasing returns industries (Arthur, 1994). In fact, existing and emerging technologies give rise to specific capabilities to new industries in a chain of networks (Arthur, 2009, Chap. 10) as briefly exhibited in Case B.

The evolutionary structure of industry development can either consist of extension of existing industries through increasing returns mechanisms and/or the formation of emerging industries through radical innovation in science and technology. This again is part of a network of private/public innovation i.e., an innovation system. The interconnection becomes clear as we develop the links in what follows in Case C.

Networks in communication or physical flows, in principle, carry the advantage of speed but they come at a price of vulnerabilities in blockage and attacks (shocks). Structural features of networks are important and guide the economies of networks. Except for physical flows topologies of networks are more important than metrics which means that distances between two points lose their significance as long as access is assured (the death of distance paradigm). This in particular applies to the Internet.

Weak vs. Strong Links. In some communication networks weak links between nodes could be more important or payoff-relevant than strong links. The sociologist Granovetter (1973) in studying the labor market showed that information from acquaintances rather than friends or relatives often proved more useful for job seekers. This knowledge could be at the very heart of expanding social networks, in particular, if they care for targeted professional links such as LinkedIn. Thus the size of a network counts and remote nodes and links could add economic value to the network.

Small World. Further, we have the ‘small world phenomena’ of Watts amd Strogatz(1998) which shows that through random graphs initial small networks expand very quickly to far distance nodes. Everything and everybody could be shortly linked.

Existence of Hubs. If we have heterogeneous networks hubs may appear as an agglomeration of central nodes. Hubs may have some immediate economic significance such as cities or corporations, transfer stations or transaction centers.

In the dynamics of networks we observe ‘clustering effects,’ i.e., ‘the rich getting richer,’ a positive feedback effect, or in economic terms of industrial competition ‘the winner takes most or all.’ The mechanism originated with physicists Barabasi and Albert (Barabasi, 2010) on mathematical models of the growth of networks. They imagined a graph starting with a small set of nodes (2 or 3) preferring old ones that already have established links. Such mechanism is called ‘preferential attachment.’ In a preferential attachment mechanism for network growth new nodes connect preferentially with old ones that already have a high degree of closeness. This evolutionary mechanism rooted in ’history and expectation’ could be one form of growth in diversity if other rules are imposed such as Schumpeter’s (1942) creative destruction or Arthur’s (1994) increasing returns mechanism, etc. A preferential rule of connectivity also applies to hubs. Hubs offer more opportunities, attractions and choices than individual nodes. The hubs are the strongest argument against the vision of a egalitarian cyberspace. The presumption is, in a scale free Internet, as in the World Wide Web, that early nodes are more likely to get connected and in forming hubs than those that appear later because the later ones have less time to link with further nodes. However, disruptions can occur, i.e., that new entrants with significant innovations who come later may have potentials to take over and grow much faster. One example was the takeover of Google in a search engine market that was already occupied by the likes of Yahoo and Alta Vista. If we view products and nodes in a complex market network and consumers as links to them then competition between them could be interpreted as a dynamic interaction of (re)assigning (that is rewiring of) links to nodes.

Fitness. In a competitive environment each node has a certain fitness. It is a quantitative measure of a node's ability to stay in front of the competition. Fitness may be related to quality of any sort. This factor emerges from the Bose-Einstein condensation in statistical mechanics (Barabasi, 2010; Newman et al., 2006). The fittest node will have the best chance to grow into the biggest hub. The fitness model allows us to describe networks as competitive systems in which nodes fight for ultimate survival on maintaining and increasing links. In networks that display fit-get-rich behavior competition leads to a scale free topology. The Bose-Einstein condensation offers also the theoretical possibility that in some systems the winner can grab all the links (winner takes all). In that case the scale free topology vanishes.

In a network such as the World Trade Web a node representing fitness has a larger GDP so that a natural linking rule is that the higher the fitness of two countries the higher the probability that they get connected. So trade links with high GDP countries are stronger and more frequent than links between low GDP countries.

Centrality of Nodes. Largeness of a country, its infrastructure, its diversity of industries and their innovation paths are major facilitators of more trade links and a boost to its growth potential. Of course, centrality patterns could change in the course of development. A hub-and-spoke network in transportation (physical and communicational) is a major facilitator to economize on transportation links and therefore strengthen competitive positions vs. rivals.

Closeness in Random Networks

Suppose a network is random in each of its nodes, their assignment subject to random mechanism, the shortest path to its farthest parts may only be a few degrees away. This characterizes a 'small world.' Suppose an individual random node is linked to 100 friends, and each of those friends have another 100 friends. Then the first node is only two steps away from 100 × 100 = 10,000 people. Again to keep up the expansion by 100, the numbers are growing exponentially by the power of 100 with each step, thus linking to 1 million, 100 million, 10 billion after 3, 4 or 5 steps.

Without digging into the mathematical analytics and ramifications the gist of the expansion shows that through connectedness we easily reach the limit of a globalized frontier after relatively few steps.

In the original scale free network topology there is no place for dominant latecomers in a market network because in such a model all nodes are identical or exchangeable that is they would not differentiate themselves in key characteristics. For example, in economic terms, they would be all identical in performance that is atomistic in the market universe.

However, in economic networks we may have asymmetric market situations as described in Case D.

Robustness of Networks. Network size and interlinkage of hubs play a role in the extent of robustness of the network. If you remove a significant number of single nodes the functioning of the network will mostly survive, however, if you do the same with big hubs then the network would be fragmented and may collapse. The distributed type structure of the Internet is a favorable topology because it can easily avoid failure nodes by circumventing them in any possible direction (Gottinger, 2014). Among the three basic network structures, centralized, decentralized and distributed only the latter one is fully able to escape vulnerabilities by building redundancies into its mesh-like architecture.

A pertinent example of the vulnerability of a network economy was shown in the unification of Germany (after 1989) where major supplier and trade links of East German companies were interrupted and structurally removed in the adaptation and transformation process (Albach, 1993). We know from network theory (Barabasi, 2010) that when the largest nodes from a regular network are removed there is a critical or 'tipping' point beyond which the network is likely to break apart. On the other hand, in a transitory and adaptive phase, new nodes might enter and let the network and their benefits expand. In economic terms, consider the world's economic trade network when more active nodes like India and China added to existing trade in the 1990s world economic growth. On the average trade links expanded for those countries even more strongly. The quantity of the links as well as their strength contribute greatly of the value of the network.

1.1. Network Size, Value and Cycles

One of the most striking economic aspects of networks is how they create externalities. Network externalities occur in both the demand and supply of the network. The textbook externality is a supply externality. For example, as a negative byproduct of a factory's production, pollution spews into the air or water. Demand externalities, on the other hand, may exist for non-network goods, but they are not usually considered important enough to merit attention. For example, economists typically do not factor demand externalities into consumers' demand functions. Many models of consumer behavior assume that the average consumer's demand for potatoes, lettuce, corn, etc. for example, are formed without any reference to how many other people are purchasing these products. Certainly the number of consumers in a given market affects demand and therefore price, but an individual's demand is independent – it does not depend directly on a product's popularity in most models. Such effects are assumed away as insignificant.

Besides the supply-side economies of scale the demand-side economies of scale are commonly seen in the communications and computer industries among others. For some goods and services, a person's demand depends on the demands of other people, or the number of other people who have purchased the same good may affect a person's demand. For example, the buyer of a telephone or fax machine would have not bought it if there were no one else who had purchased or would have purchased it. When more people have purchased it the more value of a telephone or fax machine the buyer would have obtained. This is a positive network externality based on an 'actual' or 'physical' network. Moreover, for some goods, such as Microsoft Office, the individual demand for that good inherently exists but enormously increases when other people buy the same good. In an actual network, products have very little or no value when alone, they generate value or more value when combined with others (example: fax machine). In a virtual network, hardware/software network products have value even if they exist alone, however, they are more valuable when there are more complementary goods, and also there will be more complementary goods when more people use the products. Application software developers are likely to write for the platform of the operating system that most people favor. Conversely, the operating system for which more application software is written is favored by more people. The operating system with a larger market share will provide a bigger market for the application programs. At the same time, the availability of a broader array of application programs will reinforce the popularity of an operating system which in turn will make investment in application programs compatible with that operating system more desirable than investment in application programs compatible with other less popular systems. As a result, the operating system with a larger installed base attracts more buyers whereas the small and later entrant with a smaller installed base with equal or even superior quality finds it difficult to compete. As more users are attracted to a network, the size of the network grows and confers greater value to the network users. Network effects directly challenge an important principle of classical economic theory, which posits decreasing (and eventually negative) returns to scale in most markets. Also this theory basically deals with increasing returns problems in case of supply side economies of scale but ignores cases of demand side economies of scale brought about by increasing value of existing users through increased demand, i.e., through network externalities. That is, network markets offer increasing returns over a large portion of the demand curve or even the entire demand curve. Markets with increasing returns imply that bigger is better and consumers deriving more value as the number of users grows. The flip side of this situation in terms of market structure is that the strong grow stronger and the weak become weaker.

Hence, network markets provide potentially fruitful returns to firms that can make their own products as standards in markets or in aftermarkets for complementary goods. This presents the possibility of substantial first-mover advantages: being the first seller in a market may confer an important strategic advantage over later entrants because a first mover’s technology may become locked in as a standard (Arthur, 1989, Katz and Shapiro, 1986). That is to say, the first technology that is introduced into the market may gain excess momentum when many early users join in anticipation of other users hopping on the bandwagon at a later date. This strong expectation is critical to network expansion (Choi, 1997). In the end consumers already belonging to an existing network will not likely switch to a new technology, even if it is better (Economides, 1996).

The switching costs associated with transferring to an incompatible but superior technology create ‘excess inertia’ to consumers. That means consumers will not adopt a new superior technology not only because of the sunk costs they have already put in but also because values from network externalities may be lost if they switch. Network effects, therefore, could stifle innovation.

In a traditional market, where network effects are negligible or non-existent, competition turns primarily upon price, quality and service considerations. In contrast, in those markets in which network effects are significant, competition plays out in other dimensions as well: particularly in strategies to establish, maintain, and control standards for the industry. The computer industry hence suggests that network effects have played an important role in shaping the market structure and the margins on which competition occurs.

Also, increasing returns raise the possibility of leveraging a monopoly power from one market to another. Because users may be reluctant to commit to any given system unless they believe it will be adopted by many others, the ‘network owner’ may engage in a variety of strategies to discourage potential buyers from buying a smaller network regardless whether or not it is superior. Strategies include expanding the system to include complementary products offering a wide variety of complementary products at very attractive prices or through bundling. At the same time, leveraging is able to raise rivals' economic costs of competing in the marketplace.

For example, in its effort to be adopted as the next generation standard, the owner of one element of a system may enter complementary markets by engaging in alliances as part of a strategy of attracting users to its network. Consequently, rival operating systems need to ensure the provision of substantial complementary products in the market, otherwise very few buyers will try its system. As a result, the follow-on improved or complementary products markets become very difficult.

Strong network effects are therefore themselves barriers to entry, even though it is sometimes unclear whether entry into the market ought to be encouraged. Since the increasing return deters the incentive of new entrants and increases the costs of new entrants such a blunting of incentives can occur if the leveraging practice is undertaken, not primarily as part of a vigorous competitive strategy, but in part to decrease the likelihood of competitor entry, so that the dominant firm will continue to be dominant in competition for the next market. This has clearly be shown for the (Japanese) telecommunications market (Gottinger and Takashima, 2000). The unlikelihood of success for new entrants will reduce the incentives of other competitors to innovate to the extent that these competitors perceive the opportunities to profit from their innovations hindered. All of this is particularly significant because markets in which there is rapid technological progress are often markets in which switching costs are high, in which users find it costly to switch to a new technology that is not fully compatible with the older technology. The result is an increase in entry barriers.

From what follows the definition of a network externality is given by the value of a network created by the number of its nodes. Also, network externalities can exist both for the supply and demand side of the economic equation. And networks can generate negative, positive or no externalities. Network externality networks are those that decrease in value when the number of nodes increases. More 'traditional' network industries fit into this category.

1.2. Perspectives on Network Externalities

We start with a useful distinction originally suggested by Economides (1996). He divides the work on network externalities into what he calls macro and micro approaches. Macro investigations assume that externalities exist and then attempt to model their consequences. Micro investigations start with market and industry structures in an attempt to derive (theoretically) the source of network externalities. The later category is largely founded on case studies. Three of those are symptomatic. David's (1985) QWERTY study, Arthur's (1989) model, and the domination of VHS in the videotape recorder market combined, spurred theoretical and empirical interest in network externalities. The gist of David's QWERTY study is that inferior technologies through network externalities may be subject to 'lock-ins.' This might apply to the keyboard QWERTY as well as to the adoption of the VHS against the Betamax standard though with specific technological advantages of Betamax over VHS. In empirical support of network externalities, Gandal (1994) finds that consumers pay a premium for spreadsheets which are compatible with Lotus 1-2-3 (a former industry standard for spreadsheets). In other words, consumers are willing to pay for the ability to share spreadsheet information and analysis easily with other computer users. Thus he concludes that there is strong empirical support for the existence of network externalities in the computer spreadsheet market. In another paper, Saloner and Shepard (1990) test for the existence of network externalities in the network of Automated Teller Machines (ATMs). Their results support existence.

1.3. HYPOTHESES ON NETWORK EXTERNALITIES

Perhaps it is not surprising that little quantitative work on network externalities has been done. Many examples of network industries embody cutting-edge technologies, given that theoretical work on network externalities is still relatively new, data collection is fragmentary and common data sets upon which to test theories are severely limited. One particular important question emerging on network externalities is the functional relationship between the size of a network (its number of nodes) and the network's value.

Three key assumptions about the relationship between network size and network value underlie most analyses of network externalities and their effects. They relate to linear, logarithmic and exponential assumptions.

The linear assumption postulates that, as networks grow, the marginal value of new nodes is constant. The logarithmic assumption postulates that, as a network grows, the marginal value of new nodes diminishes. Network externalities at the limit in this formulation must be either negative, zero or of inconsequential magnitude in comparison to quantity effects on prices. In contrast, Katz and Shapiro (1986) make their assumptions explicit: network externalities are positive but diminish with development, at the limit they are zero. In any case, network effects diminish in importance in these models as a network grows. The third key assumption about the relationship between network size and value is the exponential assumption which in the popular business and technology press has been named 'Metcalfe's Law.' It embodies the idea of positive network externalities whose marginal value increases with network size. Robert Metcalfe (1995) states the 'law' in this way: "In a network of N users, each sees a value proportional to the N-1 others, so the total value of the network grows as N(N-1), or as N squared for large N." The validity of Metcalfe's Law is crucial to the 'increasing returns' debate on the New Economy, facilitated by the aggregation of positive network externalities in high-tech industries. One could also consider a mixture of hypotheses such as a combination of Metcalfe's Law and the logarithmic assumption, that is early additions to the network add exponentially to the value of a network, yet later additions diminish in their marginal value. The result looks like an S curve, in a calibration of network value to the number of nodes. It is based on the idea that early additions to a network are extremely valuable, but at some point 'network saturation' should take place and marginal value should fall.

In summary, the industry and hence aggregate (growth) benefits can be classified as follows:

(i) industries that show an exponential growth (through strong complementarity),
(ii) industries that show linear growth (additive benefits),
(iii) industries that show a log relationship (stable benefits).

The mixtures of those economies create the features of the new economy. Such an economy is not immune to economic cycles, but to the extent that the network economy snowballs in an upswing period, by the same token it might also along the supply chain contract faster in a downswing period but with a better chance to stabilize quicker.

1.4. Technology Adoption, Network Industries and Network Effects

We look at the main hypotheses as how they are likely to affect the adoption process of particular network industries. The linear hypothesis is the assumption of Arthur's (1989) model subject to simulation. Given a very large number of trials, technology adoption leads (almost surely) to lock-ins. Given two technologies, A and B, further R and S agents that make adoption decisions, respectively, in Arthur's model each trial represents a random walk of an ever increasing number of R and S agent decisions. As the number of trials increases, with symmetries in both technologies A and B, the split between A and B adoptions approach fifty-fifty. That is, either one of them will be adopted, and non-adoption will be most unlikely. In Arthur's analytical model, as the number of iteration goes to infinity, the possibility of non-adoption disappears.

Correspondingly, the average adoption time until lock-in will increase with decreasing probability (of non- adoption), in conformity with the linear hypothesis, in other words, more agents become (linearly) more convinced to adopt either way. This suggests that the network effect leaves only a neutral impact on the innovation process. Against this benchmark, when the value of network size grows logarithmically in relation to its size, the average time until lock-in occurs is extended. What appears surprising is how much the logarithmic assumption delays lock-in. That is, the logarithmic specification creates less growth prospects and greater instability by delaying (or preventing) adoption from occurring. In contrast to the logarithmic hypothesis, the exponential assumption shortens the average time until adoption occurs. The average adoption is affected just as drastically by the exponential assumption as by the logarithmic one. With the exponential assumption, however, the average adoption occurs much earlier than in the baseline case. No wonder, that on an aggregate scale across network industries, it is this network effect that lends support to 'increasing returns' by the proponents of the New Economy. It can even be reinforced by speed of transactions, for example, enabled through large scale broadband internet technologies. This would support a scenario of a sustained realization of an exponential assumption as even more likely. If it can be established that the Internet triggers a technology adoption process in the form of a large and broad wave ('tsunami') across key industries, sectors, regions and countries, then increasing returns will generate exceptional growth rates for many years to come. For this to happen there should be a critical mass of network industries being established in an economy. Then an innovation driven network economy feeds on itself with endogeneous growth. It remains to be determined, empirically, which mix of sectors, with network effects with exponential, linear and logarithmic relationships will have a sustained endogeneous growth cycle.

From a slightly different perspective, it is interesting to note that the logarithmic assumption creates instability in the models. Metcalfe's law, on the other hand, which leads to immediate adoption, creating a dynamics of its own, would prevent many contemporary models from reaching equilibrium.

1.5. NETWORKED INDUSTRIAL ORGANIZATION

The development of the Internet and its use for business transactions, B to B or B to C, would make a good subject for the analysis of business cycles for several reasons. First, the Internet might constitute a formal mapping of business transactions for a networked economy, or major parts of their industries. Second, the Internet itself can be modelled and designed as a set of transactions among various agents, consumers, network suppliers and services, that reflect the size and complexity of a distributed computing system (Gottinger, 2014). That is, the Internet provides prototypical examples of a positive network externality industry. Unfortunately, the Internet is ‘too recent’ to make it an eligible candidate for a statistical study on the value of a network. As a proxy we may use the US telecommunications network over the past 50 years. There are some intuitive reasons. In many industry studies, where positive network externalities are defined, Economides (1996), Spulber and Yoo (2009), telecommunications is named as the signature network industry. The telecommunications industry is somehow described as a ‘typical’ network industry that makes it a logical place to begin a search for empirical evidence. Due to its structural similarity with other network industries like railroads, airlines, and the Internet, conclusions reached about network externalities in the communications system are arguably applicable to all of the rest.

In correspondence to the classification (i) to (iii) we would venture the hypothesis that (i) exponential growth would likely be associated with an emerging, broadly based advanced technological, strongly growing network industry, (ii) a linear relationship would be tantamount to a maturing, structurally stable industry, while (iii) a logarithmic shape would go with a technologically mature, well established industry.

The weighting of factors (i) to (iii) would characterize the scope and degree of a new economy, the higher the share of (i) and possibly (ii) the stronger the scope of a New Economy though a sizable share of (i) to (iii) would form the basis of a New Economy.

We conjecture that the higher (i) and (ii) in a sizable share of (i) to (iii) the stronger the cyclicity of the economy and the higher the volatility of movements.

1.6. CASES

Case A. *The Internet as a Reflection of the Economy*

The economic model consists of the following players: Agents and Network Suppliers. Consumers or user classes: Consumers (or user classes) request for quality adjusted services on demand or quality of services (QoS). Each user class has several sessions (or user sessions). Users within a class have common preferences. User classes have QoS preferences such as preferences over packet-loss probability, max/average delay and throughput. Users within a class share resources.

Agents and Network Suppliers: Each user class is represented by an agent. Each agent negotiates and buys services (resource units) from one or more suppliers. Agents demand for resources in order to meet the QoS needs of the user classes. Network providers have technology to partition and allocate resources (bandwidth and buffer) to the competing agents. In this competitive setting network providers (suppliers) compete for profit maximization.

Multiple Agent-Network Supplier Interaction: Agents present demands to the network suppliers. The demands are based on their wealth and QoS preferences of their class. The demand by each agent is computed via utility functions which represent QoS needs of the user classes. Agents negotiate with suppliers to determine the prices. The negotiation process is iterative, where prices are adjusted to clear the market; supply equals the demand. Price negotiation could be done periodically or depending on changes in demand.

Each agent in the network is allocated a certain amount of buffer space and link capacity. The buffer is used by the agent for queueing packets sent by the users of the class. A simple first-in-first-out (FIFO) queueing model is used for each class. The users within a class share buffer and link resources.

Agent and supplier optimality: Agents compete for resources by presenting demand to the supplier. The agents, given the current market price, compute the affordable allocations of resources (assume agents have limited wealth or budget). The demand from each agent is presented to the supplier. The supplier adjusts the market prices to ensure demand equals supply.

The main features from the economic model are:

- Characterization of class QoS preferences and traffic parameters via utility functions, and computation of demand sets given the agent wealth and the utility function.
- Existence and computation of Pareto optimal allocations for QoS provisioning, given the agent utility functions.
- Computation of equilibrium price by the supplier based on agent demands, and conditions under which price equilibrium exists. Price negotiation mechanisms between the agents and suppliers.

Case B *Increasing Returns Mechanism*

An increasing returns (IR) economy is a natural extension of a network economy induced by technological changes applied to newly emerging industries. Though not making recourse to the network economy more recent work on IR, e.g., Quinzii (1992), shows the extension of economies of scale, marginal cost pricing and efficiency to the main body of economic theory.

Most industrial sectors of highly industrialized economies are not perfectly competitive. They are usually formed by a small number of big firms with non-negligible market share. Besides being prevalent in the economy, big firms cluster around concentrated industrial structures which exhibit a skewed distribution of firm size and market share. This situation may be brought about by the intrinsic potential of dynamic technological competition to end up in (temporary) technological monopoly, so in those cases industrial competition may start out symmetric but end up asymmetric.

In IRM, Chapter 4, we show how the competitive process proliferates in increasing returns industries (IRIs) where the total of all unit activities linked together yield a higher return than the sum of the individual unit activities operating separately. For this to be happening we must show that a variety of increasing returns mechanisms combine to enable the effect of an increasing returns industry.

We propose an integrated framework going beyond previous attempts by W. Brian Arthur (1994) to provide tools and insights for explaining competition among skewed industrial structures. However, this is only a tentative step toward attempting to explain the path-dependent, indeterminate, suboptimal, locking-in nature of technological competition under increasing returns. We partially review the literature on the dynamics of technological diffusion, substitution, and competition. The purpose of this review is to show that we cannot accurately understand industrial competition without taking into account the self-reinforcing nature of commercial success in most emerging markets, a repetitive theme since Arthur's stimulating piece on 'Complexity and the Economy' and the recent follow-up book (Arthur, 1999, 2015). We enrich increasing returns mechanisms by incorporating a set of stronger, yet neglected, increasing returns mechanisms, i.e., reputation effects, infrastructure effects and positive network externalities into a preliminary framework model. The resulting theoretical structure, we will argue, captures the interdependent and cumulative character of the three aspects of industrial competition: the number and size of firms, skewed industrial structures, and the nature of technological competition.

The increasing returns discussion in economics has provided important insights into the characteristics and dynamics of modern industrial economies. However, the discussion on policy applications has (mis)led some authors and policy analysts to conclude that a completely *new economy* is emerging and that it obeys a set of rules, which are totally different from those that apply to traditional sectors of the economy. While it is undeniable that the increasing returns paradigms remain fairly new and revolutionary and while there is no doubt that this paradigm is key to our understanding of new industrial sectors, and their sustaining role in productivity growth, we should clarify its proper role in industrial structure and growth of the economy. Arthur's work generally shows that if there are increasing returns, then the processes of capital formation and labor force growth result eventually in a steady exponential growth in per capita income. Such a process could be equally well applied to corporate, industrial and economic growth of a national economy. At this stage we are most concerned about the catalytic role of technological competition in increasing returns industries. Increasing returns industries are nowadays most likely to be identified with high technology industries, in particular with information, communication and health care related industries (Gottinger, 2003).

For those industries Shapiro and Varian (1999) have suggested a combination of supply-side scale economies and demand-side scale economies to explain the intrinsic aspects of technological competition. It appears however that this way of seeing technological competition is too simple to capture the variety and complexity of real-world businesses in those industries Thus we suggest a general framework to describe technological competition in what we are going to call the *increasing returns economy*.

Case C. *Open Source Technologies (OST)*

A special feature of the IT industry is its operation through open source production systems (Evans et al., 2006). Lerner (2012) considers the rise of open source of pertinent industries (such as software) as an extreme form of decentralizing the innovation process.

Suppose there is a contestable market with network externalities engaging an incumbent and an entrant. The incumbent, unlike the entrant, already has an installed base of consumers.

We look at decision situations of firms regarding how proprietary they want to make their technology, either through patent protection or through development in open source systems (OSS). In Gottinger (2006) we explicitly modelled the direct and indirect effects of network externalities. For example, more software companies are willing to produce programs for an operating system (OS) if it has a larger consumer base. This increased competition could lead to an improvement of the quality of the OS. The model predicts that using open source technologies is likely to enhance the rate of R&D, and consequently the quality of the product. An incumbent that would choose this strategy is likely to deter entrance of a newcomer because it can play out its advantage of a larger network.

Cooperative R&D is a crucial phenomenon both from the point of view of the individual firm, its supply chain and the economy as a whole. Since innovation could be regarded as a public good, society as a whole benefits from innovation. However, the private benefits to a firm from innovating are likely to be different from the social benefits. In the absence of any mechanism preventing it, the benefits to an innovating firm are likely to be quickly dissipated by the entry of other imitating firms. In such a scenario, firms are unlikely to innovate. Thus according to conventional thinking, firms need to have some sort of reward for innovating. Intellectual property rights such as patents and copyrights provide this compensation. A big portion of the R&D literature has focused on the optimal patents' duration and breadth and the incentive of firms to innovate.

However, a different trend has emerged these days especially with the increasing proliferation of high technology network industries. Instead of trying to get exclusive ownership rights, an increasing number of firms are making their technology freely available i.e., their technology is no longer proprietary, i.e., 'open source.'

In the so-called browser war, in the 1990s, we have witnessed intense competition in the market for internet browsers between Microsoft and Netscape. Netscape had a major head start on Microsoft, controlling 90 percent of the browser market by 1996 before Microsoft started aggressively selling in the market. With the entry of Microsoft, both firms engaged in a race to have the best available product.

Given the intense competition between the two firms, by the end of 1997 Microsoft was pricing the Internet Explorer free. In contrast, Netscape was charging corporations licensing fees for using their browser. By the end of 1997, Microsoft had stolen a large chunk of Netscape's market share. Netscape eventually followed suit and started giving away its browser free. The extended battle between Microsoft and Netscape had its toll on the profits of both companies. In 1998 Netscape came up with a new strategy and decided to release its source code, the actual line of programming language, for the Netscape Communicator. This allowed users and developers to look inside the workings of the browser, to modify the software and even to redistribute the new version under their own brand name, provided that the modified source code was also freely available. The whole idea is to turn the entire internet community into a vast research division for Netscape browser.

The term 'open source software' has been widely used in the popular and professional literature. Instead of keeping their technology proprietary, the firms will distribute it freely at a price.

Even though the whole unorthodox open approach may seem counterintuitive, Netscape was not the only one who employed it. It got steam in other industries. Apache, a program for serving world wide web sites, and Sendmail, a program that routes and delivers internet electronic mail, are examples of free open source programs that dominate the market. Open source approaches have been expanded to the biotechnology and health care industries. Linux, an increasingly popular operating system created in 1991 is another classic example of successful open source software. Many of the programmers and software designers advocating OSS may share a utopian vision of software development, or they may simply want to prove themselves to be better than software giant Microsoft. However, the whole idea of OSS may not be so anti-capitalistic as it seems. It is hard to believe that profit-aiming firms will employ the OSS strategy without considering more pragmatic matters. As recently being pointed out, 'open source software' such as Linux refers to freedom of choice not free prices. The emergence of OSS as an observable phenomenon may be because the markets under consideration are no longer conventional markets. These markets exhibit 'network externalities' – a market has network externalities when buyers of a good exert positive benefits on the other users of the same good. For instance consumers are likely to value computer hardware more the more users of the hardware there are. This could be because there is likely to be a better support system the larger is the network of consumers buying the product. Similarly, it is more likely for improved software to get written for the computer hardware the bigger is the network of consumers buying it. But network externalities can be working both ways, positive or negative, for example, incompatibility with network systems on a rival operating system (such as MS Office) is a major obstacle in the OSS pursuit of the desktop though low prices for OSS products would be a strong incentive to switch and for new customers to enter.

In such a market time is of utmost importance in the race for product improvement. Firms cannot afford to let their competitors get ahead in the race for technological innovation since that would give them the added advantage of a bigger network. Also, consumers in these markets tend to exhibit a very high level of loyalty. That is because learning to use the product involves a cost. Once a consumer becomes familiar with particular software, she is unlikely to switch to a completely different brand performing the same tasks. Instead, she would rather purchase new releases of the same brand even though there can be various close substitutes with similar qualities available in the market. This enhances the effect of network externalities in the long run. Further, OSS *can* feasibly translate into better quality in markets such as those for computers.

The effect of OSS on product improvement is two-pronged. Making the technology freely available means that there can be more people directly working on improving the product. For example, ever since Linux went fully OSS, thousands of programmers have volunteered elaborate improvements of their own design for no more reward than the respect of the geek subculture. It is like expanding the R&D department, so larger improvements in quality can be realized. Second, there is likely to be a better supply of complementary goods. For instance giving out the source code for an operating system is likely to lead to more software being developed for it, which is in essence equivalent to having a better quality operating system i.e., consumers now find this more attractive.

On the other hand, making technology freely available means a loss in license fees. There is also the fear of technology being stolen. But in a market with network externalities, if the firm giving away its technology already has a sufficiently big network then it is more difficult for other firms just entering the market to steal the technology and get ahead since they would also have to overcome the network advantage of the existing firm (Gottinger, 2003). Besides in this digital era, the relative ease of creating software with similar functionalities using different programming codes has made the whole idea of keeping technology proprietary less relevant.

Case D. *Asymmetric Competition in Network Markets*

Many high technology industries are characterized by positive network externalities. Firms essentially compete and cooperate on R&D and the production of goods and services that share a network. Some models of competition contain special features that apply equally well to network markets. One is the uncertainty in technological development or uncertainty in the realization of a firm's R&D effort. The other is the dynamic nature of price competition between firms in the presence of network effects.

Firms compete with each other over an extended period of time and must therefore strategically choose prices as the market shares of the firms evolve.

Most existing models of network markets focus on only one of the two features. Further, almost all the models have the commonality that one firm captures the market instantaneously and sells to all consumers from then on (the 'winner-takes-all' market). However, the history of high technology industry abounds with instances where rival firms (and technologies) have had extended battles for providing the industry standard (Uttenback, 1994).

In network markets evolution of market share is a very interesting phenomenon especially in the face of uncertainty about future product quality. In Gottinger (2009) we attempt to provide a model framework that could capture the richness of market share evolution in the presence of network externalities.

We are looking into the following situation considered by Evans and Schmalensee (2001, 10-12)

> "Firms that are not leaders in network industries generally have little hope of reaching that status unless they come up with a major innovation – one that can defeat the natural advantage that network effects bestow on the industry leaders. Incremental innovation – making slight improvements in the leaders' products – will not enable a small firm to overtake a leader that enjoys the benefits of network economies. …It is not atypical for a fringe firm that invests heavily to displace the leader by leapfrogging the leader's technology …"

This situation described may also be encountered in 'dynamic oligopoly' where exogeneously emerging new technologies are rapidly eroding costs or where market structure responds endogeneously to intense racing behaviour (Shapiro,1989, Sec.5; Baumol, 2002, Chapters 1, 4).

We are encountering the problem of tradeoff between 'network dominance' and 'radical innovation' that could tip the market the other way, with a significant caveat added that breakthrough R&D is highly uncertain. From a strategic perspective, in this environment, for any two firms of asymmetric size, both compete dynamically over prices to win market share. In this dynamic process there are two ways to achieve (temporary) monopolistic status. The 'smaller' firm can use dynamic pricing competition to delay the time in which the 'larger' firm wins a critical market share in the hope to hit the innovation first and displace it. If the innovator is 'patient' and the probability of innovation and the discount factor are sufficiently high, there is an equilibrium in which duopoly persists (no firm achieves a critical market share) until one of the two firms wins the race for innovation.

We briefly summarize a historical example from markets for IT products. In the personal computer (PC) operating system market. Microsoft products (MS DOS and Windows) have been dominating the market since the mid 1980s. An operating system is the fundamental program that controls the allocation and use of computer resources. Thus, the utility that operating systems provide to consumers depends on the number of compatible applications. As a general rule an application that relies on a specific operating system will not function on another operating system unless it is ported to that specific operating system. Therefore, because of its dominance, the majority of applications have been written to run on Microsoft operating systems (MS DOS). The domination of MS DOS had become even stronger since the arrival of Windows 95 in the PC operating system market. This, in turn, has provided a great indirect positive network externality to PC owners who adopted Windows 95 as an operating system. Many other firms, such as IBM and BEA Systems, Inc., introduced their own operating systems and tried to compete with Windows. These products, however, lacked sufficient compatible applications to efficiently compete with Microsoft products. The lack of compatible applications prevented enough application developers and consumers from regarding OS/2 Warp or BeOS as a viable alternative to the dominant incumbent Windows. This obstacle prevented these potential entrants from obtaining a sizable market share. Their failure to enter the market successfully, however, was not due to the inferior quality of their operating systems. In fact, OS/2 Warp was reported to be at least as good as Windows, and BeOS offers superior support for multimedia applications and systems security. If consumers who use multimedia applications frequently adopt Windows at the expense of BeOS, they have to give up the convenience that is provided by BeOS. Thus, for multimedia specific users, adopting BeOS as their operating system at the expense of another operating system might provide the highest utility. Nevertheless, the lack of compatible applications, which in turn implies the lack of positive network externality, has prevented consumers from adopting OS/2 Warp or BeOS. As a conclusion, the positive network externality for the dominant incumbent (Windows) has worked as an entry barrier against entrants (OS/2 Warp or BeOS) which do not have network externality. Such an entry barrier could have only been overcome by a radical innovation, virtually leapfrogging the dominant incumbent.

REFERENCES

Albach, H. (1993). Zerrissene Netze, Eine Netzwerkanalyse des ostdeutschen Transformationsprozesses (Torn Networks, A Network Analysis of the East German Transformation Process), Berlin: Ed. Sigma.

Arthur, W. (1989). “Competing Technologies, Increasing Returns and Lock-in by Historical Events,” *Economic Journal*, *99*, 116-131.

Arthur, W. B. (1994). Increasing Returns and Path Dependence in the Economy, Ann Arbor: The Univ. of Michigan Press.

Arthur, W. B. (1999). “Complexity and the Economy,” *Science, 284*, 107-109, April 2.

Arthur, W. B. (2009). The Nature of Technology, New York: Free Press 2009.

Arthur, W. B. (2015). Complexity and the Economy, Oxford: Oxford Univ. Press.

Barabasi, A. L. (2010). Linked, The New Science of Networks, Cambridge, Ma: Perseus Publ.

Baumol, W. J. (2002). The Free Market Innovation Machine, Princeton, N. J.: Princeton Univ. Press.

Besanko, D., Dranove. D. & Shanley, M. (2000). Economics of Strategy, 2nd edition, New York: John Wiley and Sons.

Choi, J. P. (1997). "Herd Behavior, the Penguin Effect, and the Suppression of Informational Diffusion: An Analysis of Informational Externalities and Payoff Interdependency," *Rand Journal of Economics*, *28*(3), 407-425.

David, P. (1985). "CLIO and the Economics of QWERTY," *American Economic Review PP*, *75*, 332-337.

Economides, N. (1996). "The Economics of Networks," *International Journal of Industrial Organization*, *14*, 673-699.

Evans, D. S. & Schmalensee, R. (2001). "Some Economic Aspects of Antitrust Analysis in Dynamically Competitive Industries," National Bureau of Economic Research, Cambridge, Ma, Working Paper 8268, http://www.nber.org/papers/w8268.

Evans, D. S., Hagiu, A. & Schmalensee, R. (2006). Invisible Engines-How Software Platforms Drive Innovation and Transform Industries, Cambridge, Ma.: MIT Press.

Gandal, N. (1994). "Hedonic Price Indexes for Spreadsheets and an Empirical Test for Network Externalities," *Rand Journal of Economics*, *25*(1), 160-170.

Gottinger, H. W. & Takashima, M. (2000). "Japanese Telecommunications and NTT Corporation: A Case in Deregulation," *International Journal of Management and Decision-Making*, *1*(1), 68-102.

Gottinger, H. W. (2003). Economies of Network Industries, London: Routledge.

Gottinger, H. W. (2006). Innovation, Technology and Hypercompetition, London: Routledge.

Gottinger, H. W. (2009). Strategic Economics of Network Industries, New York: Nova Science.

Gottinger, H. W. (2014). "Internet Economics of Distributed Systems," *Transactions on Networks and Communication (TNC)*, *2*(6), 57-72.

Metcalfe, R. (1995). "Metcalfe`s Law: Network becomes more valuable as it reaches more users," *InfoWorld*, *17*(40), 53.

Granovetter, M. S. (1973). "The Strength of Weak Ties," *American Journal of Sociology*, *78*, 1360-1380.

Katz, M. & Shapiro, C. (1986). "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy*, *94*, 822-841.

Lerner, J. (2012). The Architecture of Innovation, The Economics of Creative Organizations, Boston: Harvard Business Review Press.

Newman, M., Barabasi, A-L. & Watts, D. J. eds. (2006). The Structure and Dynamics of Networks, Princeton, N. J. and Oxford: Princeton Univ. Press.

Quinzii, M. (1992). Increasing Returns and Efficiency, New York, Oxford: Oxford Univ. Press.

Saloner, G. & Shepard, R. (1995). "An Empirical Examination of Adoption of ATM," *Rand Journal of Economics*, *26*(3), 479-501.

Schumpeter, J. (1942). Capitalism, Socialism and Democracy, New York: Harper.

Shapiro, C. (1989). "Theories of Oligopoly Behavior," *Handbook of Industrial Organization*, Vol. *1*, Amsterdam: North Holland, 329-414.

Shapiro, C. & Varian, H. (1999). Information Rules – A Strategic Guide to the Network Economy, Boston: Harvard Business School Press.

Spulber, D. F. & Yoo, C. S. (2009). Networks in Telecommunications, Economics and Law, Cambridge: Cambridge Univ. Press.

Uttenback, J. M. (1994). Mastering the Dynamics of Innovation, Boston:Harvard Business School Press.

Watts, D. J. & St. H. Strogatz (1998). “Collective Dynamics of ‘Small World’ Networks,” *Nature*, *393*, June 4. in Newman et al. (2006), 301-303.

Chapter 2

NETWORKS, TECHNOLOGY AND COMPETITION

In traditional industries the economic model conventionally used to benchmark the market function is perfect competition. Perfect competition theory assumes that individual economic agents have no market power. The agents in the economy are price takers. It is generally believed that competition will drive the market price down to the competitive level (equal to the marginal costs) and consumer welfare is improved through allocative and production efficiencies. The perfect competition or nearly perfect market is based on the assumption in a market containing many equally efficient firms, each firm faces a perfectly horizontal demand curve for a homogeneous product, and that firms freely enter or exit the industry.

It is against this benchmark that network industries are a world apart. In their markets products are heterogeneous, differentiation in products is common, the life cycles of products are short, sunk cost is significant, innovation is essential and sometimes ‘Only The Paranoid Survive’ (A. Grove, 1996) due to technological racing. In some industries only a handful of participants are in the market and the dominant firms may easily raise the barriers of market entry to exclude competitors. In other words, in network industries, markets usually involve enormous capital and highly risky investment, economies of scale, intensive and interdependent technologies owned by different market players, network externalities of products, and a tendency of product standardization. In view of this, market failures in those industries appear significant.

In this chapter we give a fairly extensive characterization of network industries and discuss some of the competitive and market issues.

2.1. NETWORKS AND NETWORK INDUSTRIES

In conventional terms empirical examples of network industries embrace electricity supply, telecommunications and railroads. Network industries can be defined as those where the firm or its product consists of many interconnected nodes, where a node is a unit of the firm or its product, and where the connections among the nodes define the character of commerce in the industry. Railroads, for example, are a network industry. The nodes of a railroad (its tracks, rolling stock, switches, depots) are scattered across a geographic area, and the configuration of the nodes determines where, when, to whom, how much, and how quickly goods can be transported. Like the railroads, the entire transportation sector can be analyzed as a network industry. Be it airlines, trucks, or ships, each mode of transport has its network of nodes. Other network industries include utilities, telecommunications, computers, information and financial services. The nodes of these industries are units like electricity lines, phone sets, personal computers, and information platforms. The number of nodes and connected consumers may grow in tandem, but a distinction exists between an additional node on the network and an additional consumer. A node is a capital addition. This distinction is of some importance for the analysis of positive, negative, or negligible externalities.

We can distinguish many diverse features of network structures as exhibited in most of the network economic literature, e.g., Easley and Kleinberg (2010), but the follow-up forms are of central relevance for network economics and mixtures of those forms appear frequently in economic applications.

Star Networks: a star network has a collection of nodes clustered around some central resource. Movement of resources or products from one node to another must always pass through this central node. A simple example of a star network is a local telephone exchange, a call from any node (phone) must be transmitted through the central switch.

The figure illustrates the central resource CR with N denoting network nodes.

Tree Networks: In a tree network, the purpose of the infrastructure is a movement between the central resource and the nodes. Flow among the nodes is not part of the system design. Examples of systems like this include the distribution of public utilities like water, electricity, and natural gas. In each of these systems, resources flow from a central area to many outlying points.

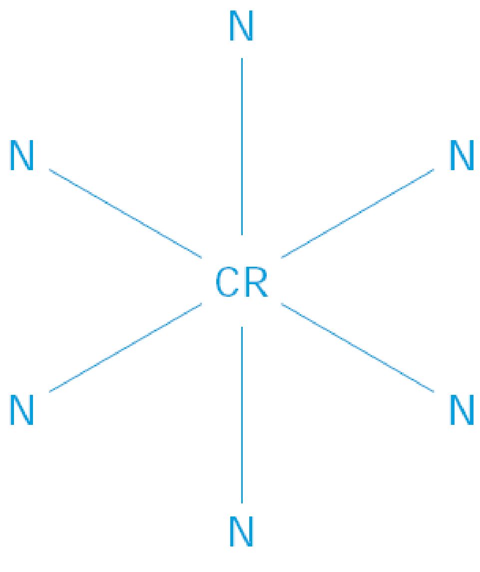

Figure 2.1. Star Network.

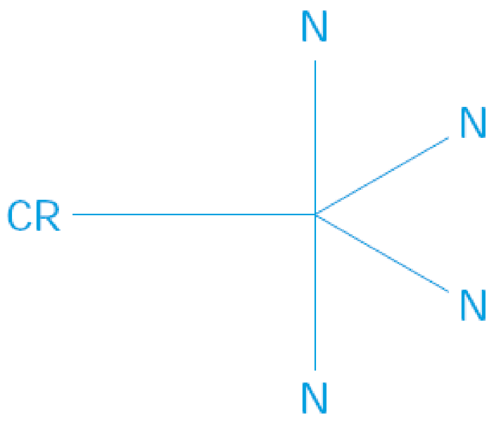

Figure 2.2. Tree Network.

Tree-structured networks can be considered one-way networks because resources flow in only one direction. On the other hand, if the system is defined broadly enough, tree-structured networks can also move resources in both directions. Consider, for example, a combined water and sewage system. Water flows out from the central resource to the nodes along a tree structure, and sewage returns from the nodes to the central receiving area through the same structure.

Crystal Networks: A crystal network occurs when the central resources are distributed among connected star networks. Like a star network, movement can occur from any point on the network to any other point. Unlike the star network, that movement will not always traverse a single central point. Movement from one star to the next will involve both central connections, while movement within a star will require only one. An example of a crystal network is the long-distance telephone network which is really the connection of many local networks through medium and long-distance lines

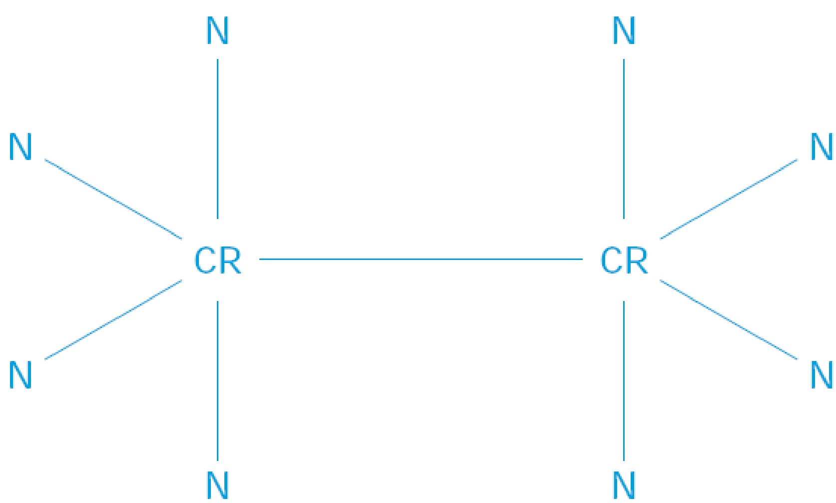

Figure 2.3. Crystal Network.

Web Networks: In a web network, each node is connected to many other nodes. Paths between the points on the network multiply as the interconnections increase, and the network assumes a distributed character, no network center can be identified. There are two salient features to a network with a web structure. First, the resources distributed via a network with a web structure are located in the nodes. This means that as the number of nodes increases, the resources to be distributed throughout the network also increase. Second, the distributed nature of a web network makes the system less vulnerable to the failure of any part. This feature was a key attraction of networked computing projects evolving into today's Internet.

Here each node is connected to four nodes around it, and a lattice or grid emerges. That is, if you pick any two points between which data should travel, the failure of any one point in the lattice (as long as it is not the origin or destination of the data) will not interrupt communication. This is not the case with the other structures presented here. In a star network, for example, failure of the central node disables the rest of the system.

Network Characteristics: Beyond the various structures of networks, there are also a few characteristics of networks that can influence their economics. Two of the most common characteristics made about networks concern their directionality and spatial character. Most networks carry their loads to and from all the nodes on the network, traffic is reciprocal. Transportation networks and the telephone system are like this, and reciprocal networks are probably what most people imagine if they think about networks. One-way networks, however, also exist. Most of the tree structured networks, electricity, water, natural gas, deliver from a central resource to homes, but not vice-versa. A communications example of a one-way network is the cellular paging system, though the introduction of two way pagers and cellular phones shows that a transformation of that network is already underway. A spatial network is one that is fixed geographically, a railroad network is a typical example. As a spatial network, the existing track layout determines, for example, who can be served by the railroad and where goods can be delivered. A non-spatial network like the Internet, is free of these geographic constraints.

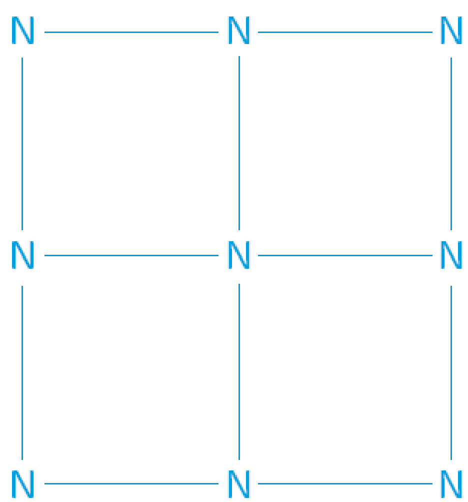

Figure 2.4. Web Structure.

2.2. High Risk Investments and Sunk Costs

A network industry company survives in the market only by maintaining rapid innovation, relying on intensive, often large scale research and development projects to develop new products or processes.

On the other hand, sizeable financial commitments receive no guarantee of profits or success, as a survey on the New Economy (Economist, 2000) appears to substantiate. Accordingly, these tremendous sunk costs make entry and exit to market relatively difficult. Such a company, therefore, inevitably requires large, irreversible investment. Further, rapid innovation shortens the product's life cycle and leads to non-price competition. The costs of new products, hence, usually increase rather than decrease. Today's IT industry illustrates this dilemma. The cost of staying at the cutting edge of technology is horrendous. Competition is relentless – unless you are sheltered from competition and heavily subsidized (Gottinger, 2013). In addition to the large investment on the supply-side, customer's investment can also be enormous and irreversible. Because today's technologies are complicated, sophisticated and through standard-setting complementary as well as compatible, customers must commit to a generation of equipment, where personal training, ancillary facilities, applications software, or other ongoing expenses are tied to the particular technologies. This means the investments of customers and their switching costs are very high (Choi, 1994).

2.3. Economies of Scale

In network industries, the fixed costs are tremendous in proportion to its total costs of the products. The returns to size of the firms are not constant. The average total cost lowers as the total output increases.

The effect of economies of scale is commonly seen in every stage. The most important assets of network industries are knowledge or information. The costs of research and development of new products or new processes are prohibitively high. Nevertheless, once it has been developed, a product (book, software) can be produced with almost zero marginal costs. This is illustrated in the computer software industry where virtually all of the costs of production are in the design of the program of software. Once it is developed, the incremental costs of reproducing software, that means, the costs of copying a disk, are negligible. In other words, the production costs of software are almost totally independent of the quantity sold and consequently, the marginal cost of production is almost zero. At the same time, unlike in other industries, the increasing returns may be of no ending. One of the characteristics of computer software industries is that there is no natural point of exhaustion at which marginal costs begin to exceed marginal revenues and at which it becomes uneconomical to increase production. That is to say, in this industry there is always the case that the more products produced the lower the costs per product. Large scale production is generally required in order to achieve minimum per unit cost rapidly when the descending learning curve, termed 'learning by doing' effects, are significant. The increasing returns to scale arise because increasing experiences foster the improvement of production technologies, and allows the managers and workers to specialize in their tasks. As a result, the unit cost fell as workers and operators learn by doing. This is particularly evident in the semiconductor industry or computer assembly industry. In those industries, very often than not, shadow costs lie below current-period costs because the presence of a learning curve allows firms to lower their unit costs tomorrow by acquiring production experiences today (chip production)

Moreover, in some industries where price discrimination is possible, the learning curve effect leads to a dynamic process whereby differential pricing increases the volume of output and expanded output results in declining costs, thereby allowing new customers being served. The presence of learning curve economies is particularly important from a public policy perspective. If there are increasing returns, then it is more economically advantageous to have one large firm producing than have many small firms.

On the other hand, the increasing returns, at the same time, handicap small-scale entry.

For example, the introduction of competition into the telecommunications market generally has two phases. The first phase allows new suppliers in the areas of value-added services. The next phase introduces competition into core areas of local, trunk (long-distance), and international services. However, in practice, high capital costs and economies of scale limit the development of competition in local services.

So far we have observed economies of scale effects from the supply side, those are reinforced by equally strong effects from the demand side which are facilitated through 'network externalities.' Also those reinforcing effects could easily lead to overwhelming market dominance and winner-takes-all situations (Shapiro and Varian, 1999).

In further expansion through platform hubs the network economy would also provide 'economies of scope' which would lead to product diversification and meet individualized product demands.

2.4. NETWORK EXTERNALITIES

One of the most striking economic aspects of networks is how they create externalities. Network externalities occur in both the demand and supply of the network. The textbook externality is a supply externality. For example, as a negative byproduct of a factory's production, pollution spews into the air or water. Demand externalities, on the other hand, may exist for non-network goods, but they are not usually considered important enough to merit attention. For example, economists typically do not factor demand externalities into consumers' demand functions. Many models of consumer behavior assume that the average consumer's demand for potatoes, lettuce, corn, etc. for example, are formed without any reference to how many other people are purchasing these products. Certainly the number of consumers in a given market affects demand and therefore price, but an individual's demand is independent – it does not depend directly on a product's popularity in most models. Such effects are assumed away as insignificant.

Besides the supply-side economies of scale the demand-side economies of scale are commonly seen in the communications and computer industries among others. For some goods and services, a person's demand depends on the demands of other people, or the number of other people who have purchased the same good may affect a person's demand. For example, the buyer of a network product would have not bought it if there were no one else who had purchased or would have purchased it. When more people have purchased it the more value of a telephone or fax machine the buyer would have obtained. This is a positive network externality based on an ‘actual’ or ‘physical’ network. Moreover, for some information goods, such as Microsoft Office, the individual demand for that good inherently exists but enormously increases when other people buy the same good. In an actual network, products have very little or no value when alone, they generate value or more value when combined with others. In a virtual network, hardware/software network products have value even if they exist alone, however, they are more valuable when there are more complementary goods, and also there will be more complementary goods when more people use the products. Application software developers are likely to write for the platform of the operating system that most people favor. Conversely, the operating system that carries more application software are favored by more people. The operating system with a larger market share will provide a bigger market for the application programs. At the same time, the availability of a broader array of application programs will reinforce the popularity of an operating system which in turn will make investment in application programs compatible with that operating system more desirable than investment in application programs compatible with other less popular systems. As a result, the operating system with a larger installed base attracts more buyers whereas the small and later entrant with a smaller installed base with equal or even superior quality finds it difficult to compete. As more users are attracted to a network, the size of the network grows and confers greater value to the network users. Network effects directly challenge an important principle of classical economic theory which posits decreasing (and eventually negative) returns to scale in most markets. Also this theory basically deals with increasing returns problems in case of supply side economies of scale but ignores cases of demand side economies of scale brought about by increasing value of existing users through increased demand, i.e., through network externalities. That is, network markets offer increasing returns over a large portion of the demand curve or even the entire demand curve. Markets with increasing returns imply that bigger is better and consumers deriving more value as the number of users grows. The flip side of this situation in terms of market structure is that the strong grow stronger and the weak become weaker.

Hence, network markets provide potentially fruitful returns to firms that can make their own products as standards in markets or in aftermarkets for complementary goods. This presents the possibility of substantial first-mover advantages: being the first seller in a market may confer an important strategic advantage over later entrants because a first mover’s technology may become locked in as a standard (Arthur, 1989, Katz and Shapiro, 1986). That is to say, the first technology that is introduced into the market may gain excess momentum when many early users join in anticipation of other users hopping on the bandwagon at a later date. This strong expectation is critical to network expansion (Choi, 1997). In the end consumers already belonging to an existing network will not likely switch to a new technology, even if it is better (Economides, 1996).

The switching costs associated with transferring to an incompatible but superior technology create ‘excess inertia’ to consumers. That means consumers will not adopt a new superior technology not only because of the sunk costs they have already put in but also because values from network externalities may be lost if they switch. Network effects, therefore, could stifle innovation.

In a traditional market, where network effects are negligible or non-existent, competition turns primarily upon price, quality and service considerations. In contrast, in those markets in which network effects are significant, competition plays out in other dimensions as well: particularly in strategies to establish, maintain, and control standards for the industry. The computer industry hence suggests that network effects have played an important role in shaping the market structure and the margins on which competition occurs.

Also, increasing returns raise the possibility of leveraging a monopoly power from one market to another. Because users may be reluctant to commit to any given system unless they believe it will be adopted by many others, the ‘network owner’ may engage in a variety of strategies to discourage potential buyers from buying a smaller network regardless whether or not it is superior. Strategies include expanding the system to include complementary products offering a wide variety of complementary products at very attractive prices or through bundling. At the same time, leveraging is able to raise rivals’ economic costs of competing in the marketplace.

For example, in its effort to be adopted as the next generation standard, the owner of one element of a system may enter complementary markets by engaging in alliances as part of a strategy of attracting users to its network. Consequently, rival operating systems need to ensure the provision of substantial complementary products in the market, otherwise very few buyers will try its system. As a result, the follow-on improved or complementary products markets become very difficult.

Strong network effects are therefore themselves barriers to entry, even though it is sometimes unclear whether entry into the market ought to be encouraged. Since the increasing return deters the incentive of new entrants and increases the costs of new entrants. Such a blunting of incentives can occur if the leveraging practice is undertaken, not primarily as part of a vigorous competitive strategy, but in part to decrease the likelihood of competitor entry, so that the dominant firm will continue to be dominant in competition for the next market. This has clearly be shown for the Japanese telecommunications market (Gottinger and Takashima, 2000). The unlikelihood of success for new entrants will reduce the incentives of other competitors to innovate to the extent that these competitors perceive that the opportunities to profit from their innovations are hindered. All of this is particularly significant because markets in which there is rapid technological progress are often markets in which switching costs are high, in which users find it costly to switch to a new technology that is not fully compatible with the older technology. The result is an increase in entry barriers.

From what follows the definition of a network externality is given by the value of a network created by the number of its nodes. Also, network externalities can exist both for the supply and demand side of the economic equation. And networks can generate negative, positive or no externalities. Network externality networks are those that decrease in value when the number of nodes increases. More ‘traditional’ network industries fit into this category.

2.5. Complementarity, Compatibility and Standardization

In network industries, many products have very little or no individual value, but produce value only when combined with other products. Since a product involves lots of technologies, or is made of different components, or a product system combines several goods and services, the demand of those technologies or intermediate goods or services are thus often interrelated. That is to say, they are strongly complementary, although they need not be consumed in fixed proportions. Those complementary products are usually described as forming systems, which refer to a collection of two or more components together with an interface that allows these components to work together. Nevertheless those components, products, or services are usually provided by different manufacturers, in fact. The products, components or services need to be compatible with each other in order to combine the components into operable systems. By the same token, other manufacturers can market their individual products only when the products are compatible with other products. This is easily illustrated in computer assembly, software industries and elsewhere. In many cases, these strongly complementary components purchased for a single system are spread over time. If the components or products of different brands are incompatible, the customers need to make their purchase decision on the total system. Besides the complementary components, a system may include the future stream of ancillary products and services over the life of a primary product. In other words, rational buyers should take into consideration availability, price and quality of the components that they would be buying in the future. As a result, customers' costs in purchasing network products are not limited to the price of the product but more importantly, also include the customer's large investment in complementary equipment or training for employees when they use the product. Likewise, whenever consumers become accustomed to the products of particular manufacturers, they do not shift to other products quickly not only because they are unfamiliar with the new products to operate with but also because the complementary equipment or system is sometimes incompatible. In short, the incumbent users switching to a new technology would lose existing network benefits and would have to replace not only the durable goods themselves, but also any sunk investment in complementary assets. It thus provides opportunities for the existing system to exploit the customers and deter competition from the rivals. Compatibility is crucial to gain the benefits in network industries, this in particular applies to the software industry (Shy, 2001). Network effects will be workable in products of different manufacturers only when their products have the same protocol or bridging technologies allowing them to communicate with each other. Even in the computer software industry, although the product will allow a single user to perform a variety of functions, whether or not others own the software, the value of a given software program grows considerably as the number of additional purchasers increase. This means that network effects will be unavailable in the software of different manufacturers, which are not compatible, unless converters exist in the interface, if the technology allows (Farrell and Saloner, 1992). Standards, especially interface standards, therefore play an important role in network industries. Without standardization networks would be confined to those users who purchased the products made by the same manufacturer if products made by different manufacturers are incompatible. The standardization of computer software hence facilitates the formation and operation of computer networks, the transfer of files among users and across applications, and savings in training costs. When standards are established, entry, competition, and innovation may be easier to handle if a competitor needs only produce a single better product, which can then hook up to the market range of complementary products, than if each innovator must develop an entire 'system.' Compatibility standards thus allow networks to grow, providing pre-competitive benefits by creating 'networks' of compatible products.

Further, because of the standard, economies of scale drives prices down, and the network becomes more attractive and grows. The benefits to society as a whole are greater when standardization allows for product compatibility among all users. (Besen and Farrell, 1994).

On the other hand, standardization may carry costs of their own. First, there may be a loss of variety, reducing the flow of services generated by the product. Second, the costs of each technology may vary between and among consumers, so that if only one technology is offered, some consumers are forced to purchase what is for them the more expensive technology. Third, and most important, is that despite their advantages for innovation, networks can also retard innovation.

Some economists argue, once standardization is achieved, it can be hard to break out of that standard and move to another, even if the original is no longer the most suitable (Farrell and Saloner, 1992). It can be hard for users or vendors to coordinate a switch from an old standard to a new one, even if all would like to do so.

Moreover, some users may have invested a considerable amount in training themselves to use the existing standard. They will be reluctant to abandon that investment even if new technology is better and a switch would be desirable for the sake of efficiency. New users must choose between the benefits of compatibility and the benefits of the new technology, and often compatibility wins the game because of the effects of 'excess inertia,' even if it should not. This is a socially undesirable failure to adopt an improved technology or facilitate improved innovation.

2.6. SETTING STANDARDS

Standardization is a natural tendency in network markets, particularly in system markets where strong positive feedback enables one system to pull away the popularity from its rivals once it has gained an initial edge. Many market cases could emerge (Weitzel, 2004). This defacto standardization generated through market competition in network industries confers on the winner the full value of the standard and not just the value of his contribution. This result is inefficient in social benefits, however. Network effects, which posit incremental benefits to existing users from network growth, suggest that network goods should be optimally priced as low as possible to allow widespread adoption of the standard. Nevertheless, just as a monopolist maximizes its revenue by raising prices above a competitive level, a single company winning and owning a propriety standard can set whatever price it wants. This is not the best way to benefit consumer welfare. A proprietary standard also seems unnecessary as it is used to encourage the production of future works of intellectual property. While the intellectual property regime posits incentives of innovation, the winner of the standards competition may receive a windfall that is far greater than what an intellectual property regime normally provides. Of course, in order to resolve the problem of de facto standardization, a central government authority may simply decree a standard. This strategy is likely not to be effective. The government organization is not market-oriented, and is very likely technologically less efficient. If they choose an inefficient standard, it will be hard to surpass. Private standard setting organizations are more efficient. If they choose the wrong standard it may be leapfrogged by a new standard. At the same time the innovating firm may believe that it can be advantageous from allowing, indeed encouraging, other firms to attach to its new product. In this case, strategic alliances including licensing and joint ventures provide a direct and explicit way to make its product standard. At the same time, if people can switch back and forth between competing systems of what essentially is the same standard, perhaps society can seize the benefits of competition without wasteful duplication of efforts and without stranding consumers who make the wrong choice. A possible solution is to make competing standards interoperable. One approach to achieving interoperable standards is for a private industry organization open to all members to adopt a single standard.

If the members of such a group collectively have a significant market share, their adoption of a standard may produce a 'tipping' effect, bringing the rest of the industry into line. That is to say, the group standard setting may trade off first-round competition (to set the de facto standard) for competition within the standard in later periods. Cooperation is therefore extremely attractive in the setting of standards in the network industry because compatible systems are almost unavoidable in this field. Private group standard-setting through strategic alliances or joint ventures will be more efficient than both by government organization and de facto standardization. Since having multiple companies participating in a standard means that those companies can compete to offer products incorporating the standard after it is selected, thus expanding output and lowering prices. Group standard setting may also promote competition in the development of improvements to the standard, since each of the competitors may seek an advantage over the others by improving the design in ways compatible with the basic interface specifications. Such group standard-setting is common in the computer and telecommunications hardware industry. It has especially been known as the computer industry reaches maturity that the open system plays an essential role. An illustrative example was the market for digital videodiscs (DVD) where two major industry groups, such as Sony-Philips and Toshiba-Matsushita, competed offering incompatible standards for several years and once agreed to use a single compatible format incorporating elements from both products. The agreement is splintered at last not because advantages of compatibility are underestimated but gross profits and possibilities of controlling next generation technology by the winning system are dominant.

Some difficulties are usually found in forming the alliances to set up a standard. The first obstacle is that, since a standard must be chosen for the long term, participants want to get it right. Thus even if all interested parties had indentical interests, there would be some delays. The second obstacle is vested interest. Unless discussions are in advance of the market, different vendors have different installed bases neither of them wants to agree on the other's standard. This bargaining problem can cause serious delay. For example, many of the tree-structured public utility networks like water and electricity distribution display negative network externalities. The key to the negative externality is that in many tree-structured networks, resources are concentrated in a central location. For the electricity grid, that location is the main power plant, for the water system, the central resource is usually a reservoir. The utility network exists to connect consumers to the central resource. When new nodes are added to the network, the number of consumers with access to the resource increases but the potential resource to be distributed remains the same. Thus adding nodes to the network divides a fixed resource among more consumers, meaning less of the resource for each user on average. One can see examples of this in developing countries when, for example, new electricity consumer demand is not met with increased power generating ability and brown out results. One can then clearly see the negative network externalities of an electrical grid. In contrast, networks may also exhibit no significant demand externalities. One may consider, for example, the many large and important broadcast networks (i.e., television, radio) that exist in the world today. Adding nodes (television sets, radio receivers) to these networks creates little or no demand externalities: the average considering purchasing a television or radio does not care what percentage of other consumers own these products. Value in these networks is created centrally, and the nature of television and radio broadcasts is such that reception by one instrument has no impact on others' reception. Finally, many of the high technology networks flourishing today embody positive externalities. Networks like these increase in value for the consumer as nodes are added and the system expands. One of the best examples of these kinds of networks is the telephone system. What makes the difference between negative or no-externality networks and positive externality systems? The key resides in the location of resources. Where the electricity grid or the water pipelines or the radio broadcasts are built around a valuable central resource, the telephone network is used to bring highly distributed resources together. When a new phone is added to the network it makes demands on the central resources of the network (switching, for example) but it also adds the resources of the newly connected people to the network. Even more striking is the example of the world wide web. New servers (nodes) on the world wide web bring with them new web users who bring new information with them. The attraction of the world wide web only grows as more people make more information available.

2.7. The Rationale of Strategic Alliances

To do business in an international or globalized context more and more network companies decide to form strategic alliances instead of other interfirm transactions such as acquisitions or transactional arrangements. As has been shown for particular industries, strategic alliances are a natural outgrowth of network economics (Gottinger et al., 2010).

Strategic alliances of that sort rely on sharing the risks, synthesizing complementary skills or assets, and attaining economies of scale and scope. More importantly, this quasi-integration arrangement can provide participants closer long-term relationships that arms length market transactions and short-term contracts cannot offer. In network industries, strategic alliances are formed to share the costs of research and development, to complement technologies when producing new products, or to combine resources when entering new markets. Strategic alliances coordinate divergent strengths of research capacities to facilitate an efficient exchange of knowledge or a pooling of skills. More often than not the introduction of a new product, process or improving the quality of the product is possible only when these insights combine together, insights which would have been otherwise closely held as trade secrets by individuals while acting alone. Diverse skills can be brought together in a manner that creates economies of scale in product or process innovation. In addition, when the minimum efficient scale of research and development is too large relative to production and marketing, sometimes it can be performed economically only if two or more firms join together.

2.8. Special Issues Relating to Network Industries

Innovation based on Schumpeterian markets are dominated by product innovation, where incumbents can be completely displaced by entrants. These markets might be seen within the computer software industry and within consumer electronics.

Network industries give rise to a number of issues that are likely to affect the strategies used by firms to innovate. The characteristic features of these industries originate from network externalities. A positive network externality arises when a good is more valuable to a user the more users adopt the same good or compatible ones. A common distinction is that between direct and indirect network externalities.

Direct externalities typically occur in a two-way (physical) network. This externality reflects the fact, for instance, that a telephone user benefits from others being connected to the same network: an additional telephone increases the number of potential communication within the system, and thus the value of membership.

Indirect externalities arise from the fact that a network of users increases the incentives to produce compatible products that are complementary with the network good.

The typical system comprises of a ‘platform’ (or hardware) and of ‘applications’ (or software) that can only be used with a platform. Compatibility refers to the possibility that software can be used with a particular platform. The network externality arises when users make their purchase over time, because in the presence of economies of scale a greater number of complementary products can be supplied at a lower price when the network grows. The existence of network externalities has a substantial impact on the way technologies and economic mechanism designs are chosen and promoted (Hurwicz, 1999).

2.9. Technology Adoption Decisions

A number of issues that arise in markets characterized by network externalities can be addressed by considering a system only, setting aside competition between different systems. The existence of direct network externalities drives a wedge between private and social incentives to join the network. A single user does not consider the benefits that would accrue to others by his joining the network and as a result the market may lead to networks of inefficient small size. In fact, the benefits that an individual would take into consideration when choosing whether to join a network depends not only on the current size of the network, but also on its expected future size.

Consumers’ expectations play a central role in driving market outcomes where there are technology decisions in markets with important network effects. An important consequence is that multiple equilibria can easily arise. For instance, if all consumers expect no one else to join the network, then its size would be zero, even if the network may be valuable for consumers; if all expect everyone to join, the network may achieve a large size.

Indirect externalities may have a different impact on the size of the network. In a system market, one consumer’s adoption decision would not affect other consumers, given the prices and varieties of products available. The externalities arise indirectly through the impact of one consumers’ adoption decision of the future variety or prices of applications.

2.10. Product Selection Decisions

Competition in network industries may take the form of competition between incompatible systems. Network externalities may lead to a de facto standardization whereby everyone uses the same system. Due to positive feedback elements, a system may become dominant once it has gained an initial edge, a phenomenon which is usually referred to as 'tipping.'

Two main factors, however, may limit tipping and sustain multiple networks:

(i) network externalities being exhausted at a smaller network size,
(ii) consumers' heterogeneity and product differentiation.

In the latter case multiple networks would reflect consumers' love for variety and there is a typical trade-off between variety and standardization.

Competition between systems may be very intense, at least before a dominant system, if any, would emerge as dominant. For instance, firms may be engaged in very intense price competition at an early stage for seeking to establish an installed base and achieve leadership.

This competition can be interpreted as firms bidding for future monopoly profits.

Competition may also be played in trying to affect consumers' expectations about the dominant system that would emerge.

An important implication of the network externalities literature is that the market may settle in an equilibrium where the dominant system is that with a lower social evaluation. Markets can exhibit 'excess inertia' and remain locked into an obsolete standard, even though a better one is available.

However, markets can also exhibit 'excess momentum' whereby the market tips inefficiently to new technologies. This could, for instance, arise when competition is between an older technology which is competitively supplied and a new technology which is sponsored. The new technology that is sponsored may have an advantage over an older technology that is more competitively supplied because the sponsor of the technology may engage in pricing below costs or other types of investments (with the hope of recouping these later once the technology is established). Nonetheless, this perceived risk is not straightforward — pricing very low from the outset may be the only way of introducing an efficient new technology, and tipping may be inevitable.

2.11. COMPATIBILITY CHOICES

Compatibility between different networks/systems is often a choice variable of firms. Two types of compatibility can be considered:

(i) horizontal compatibility: between two comparable rival systems; and
(ii) vertical compatibility: between successive generations of a technology.

Compatibility entails both social benefits and costs:

compatibility expands the size of both networks thereby avoiding the cost of participating to two different networks (e.g., duplicate equipment); compatibility in systems may lead to lower production costs for economics of scale, learning effects, etc. compatibility enhances variety by allowing consumers to combine components from various systems; the risks of adopting a particular technology are lower; and costs derive from the mechanism by which compatibility is achieved. Standardization may lead to a loss of variety and may prevent the development of new incompatible systems; adapters that allow interfacing have a cost themselves and may work imperfectly.

The nature of competition in the market is affected by compatibility decisions. For systems that are compatible, competition is essentially at the level of each component. For incompatible systems, competition is at the network or system level.

Katz and Shapiro (1986) consider the impact of compatibility on pricing competition over time. Price competition is relaxed at earlier stages of the product life cycle because the market loses its winner-takes-all feature. For the same reason, however, competition may be intensified in later stages of the product life cycle.

Compatibility decisions may be affected by a number of factors:

(i) asymmetries in the probability of being the winner, e.g., reputation,
(ii) product differentiation and installed base.

Side payments may help firms to reach an agreement on compatibility. When they are not feasible, it is useful to distinguish markets where a firm can unilaterally impose compatibility and those where a firm can unilaterally impose incompatibility.

In general, there may be different mechanisms and institutions that govern compatibility choices: standards committees, unilateral action and hybrid mechanisms (Farrell and Saloner, 1988).

APPENDIX

As a specific example of a technology advanced network industry the telecoms bear certain structural features that link networks to competition. A seminal study on the detailed issues of this industry is Laffont and Tirole (2001).

Here we look at a shortlist of issues connected with the dynamics of the telecommunications industry that were subject to an ongoing competition policy.

(i) Regulation

Regulation should not stand as an end itself but should be used to support market forces, increase the chances of fair competition., and achieve wider social, economic and general policy objectives,

A reliance on economic incentives suggests that one should place greater reliance on the ability of market forces to ensure regulatory objectives. If this cannot be achieved in particular instances one should seek a balance between competition rules and sector-specific regulation.

Market Entry and Access. A particular issue is market entry in a given market structure. In this context, one might argue that where any network can potentially carry any service, public authorities should ensure that regulation does not stop this happening. Artificial restrictions on the use of networks, or to maintain monopolies where other parts are fully open to competition, may deny users access to innovative services, and create unjustified discrimination. Such an approach could be seen as running counter to the technological and market trends identified with convergence.

Types of restrictions are particularly important where competition is at an early stage or where a particular player enjoys a very strong position (for example, over a competing network). In such cases, specific safeguards can ensure that potential competitors are not discriminated against or that there are adequate incentives for them to enter the market. According to this argument, appropriate safeguards might take the form of accounting separation or transparency requirements, structural separation or even full line-of-business restrictions.

Access at either end of the transmission network will be of crucial importance.

In general, the terms on which access is granted to networks, to conditional access systems, or to specific content is a matter for commercial agreement between market actors. Competition rules will continue to play a central role in resolving problems which may arise. The emerging market will consist of players of very different sizes, but as indicated above there will also be strong vertically-integrated operators from the telecommunications, audiovisual (principally broadcasting) and IT/software industries building on their traditional strengths and financial resources. Issues which could arise across the different sectors include bundling of content and services, or of network capacity and services, predatory pricing, cross-subsidisation of services or equipment, and discrimination in favour of own activities.

Within the telecommunications sector, the development of the Internet is raising a range of issues connected to the terms on which Internet access providers get access to current fixed and mobile networks. One issue is whether they should enjoy the same interconnection rights as other players and whether they should be able to get access to unbundled service elements, whilst another issue is whether such providers in offering a range of telecommunications services should share some of the obligations of providing telecoms services.

Frequency Spectrum The provision of services (and the development of effective competition) will depend on the availability of sufficient network capacity, which for many services means access to radio spectrum. The parallel expansion of television broadcasting, mobile multimedia and voice applications, and the use of wireless technologies within fixed networks will lead to a significant growth in demand. The take up of wireless local loops and the arrival of Universal Mobile Telecommunications Services (UMTS) early in this century all point to a steady growth in demand for spectrum. Given the importance of spectrum, variations identified in between sectors with regard to how much spectrum is available and how much that spectrum will cost may have an important impact on the development of existing and new delivery channels. Though overall allocations are determined at an international and regional level, current differences across sectors to the pricing of frequency may create potential competitive distortions. One example could be where a broadcaster offering multimedia or on-line services uses spectrum obtained free or at low cost, competes with operators from the telecommunications sector who have paid a price reflecting the commercial value of the resource allocated.

From an economic standpoint, pricing spectrum may encourage its more efficient use and may help to ensure that frequency is allocated to the areas where it is most needed. Clearly, such an allocation also bears a lot of risk for the telecom operators. Frequency auctioning is favoured by many economists as the way to best ensure outcomes which are in the consumer's ultimate interest. Although others express concern about the impact of such pricing on prices charged to users.

(ii) The Internet

Two different developments in Internet service provision could be observed in an international context. One, mostly an American phenomenon, was an entrepreneurship drive to establish new companies to do e-business from the bottom up, while in Europe a top down approach emerged through established telecom carriers to launch internet service subsidiaries. Those spinoffs are now by far the largest ISPs, and most of the smaller players rely on the services offered by the bigger players (depending on them for the supply of internet connectivity). Given this heterogeneous market structure, network effects could bias competition to the effect that if one network becomes larger through positive feedback by comparison with his competitors, those competitors may no longer be able to compete, because their subscriber base is too small to attract customers away from the dominant network.. It is unable to offer the same access to subscribers. And although that may not matter when interconnection is free, any operator who becomes dominant may not be able to resist the temptation to charge others for connection to his network, thus further hampering competitors from offering any effective constraint to the dominant player’s pricing.

Still at present the Internet is a relatively open and competitive environment, so there is a comparatively conservative task in applying competition rules to ensure that it stays that way. But the speed of development of the Internet is such that competition issues could be raised very quickly, and therefore one needs to be vigilant when examining the sector in the context of competition law.

Internet regulation and telecoms regulation come from completely opposite directions. Telecoms regulation has been born out of the liberalization of the monopoly, single provider environment and new entrants have been introduced into the market. Internet regulation before now has been largely self-regulation, if any regulation existed at all. Now the Internet is becoming a system over which more and more business is being done, as opposed to the simple exchange and sharing of data which it used to be. This commercialization asks for a more robust regulatory system to protect users and suppliers. An important element of that protection is the assurance of the application of competition law.

We see no problem with the Internet continuing to have strong elements of self-regulation in the future. That is one of the reasons why it has developed in such a dynamic manner in the past. One has to be wary, however, that self-regulation does not lead to private monopolistic practices that run counter to competition law.

(iii) Market Access

Access (and the related issue of network interconnection) is one of the central issues in the telecommunications/media/information technology market and the way in which competition rules are applied to the players within it.

The central problem is that, given the evolving market structure, the converging sectors depend upon ensuring access to bottleneck facilities. These are essential for entering the market to reach customers. In particular, through vertical mergers and joint ventures there is the potential danger that thresholds are exceeded at which point the concentration of market power in the whole value chain - content, distribution, cable - becomes unacceptable. This can be shown in a number of recent telecom and media cases, In these cases particular developments deserve attention such as, for example, attempts by market players to gain control: for digital TV based services in view of set top boxes or decoders; or for Internet services in view of the browser, the server, and the access provider.

These are developments which link up critical elements of the future information infrastructure with other dominant market positions., that is, link ups between telecoms, online providers, and content providers.

Backbone Access

The Internet is a network of interconnected networks. In order to provide a full Internet service any operator will need access to all, or at least the vast majority, of the networks connected to the Internet. Market definition in this area is difficult: the distinction between competitors to whom an operator will provide reciprocal access to its customers, and customers to whom an operator will provide access to all other Internet users appears more fluid than is the case in traditional telecommunications, such as voice telephony.

Notwithstanding these difficulties of market definition and quantification, similar concerns in relation to the power of particular networks appear to arise as with traditional telephony and interconnection. These concerns include the risks that a dominant network operator: charges supra-competitive fees for network access; seeking to reinforce its position, for example, by concluding lengthy exclusive arrangements with its customers; or favoring its own operations at the expense of third parties.

(iv) Local Loop

Given the commercial and technological constraints on providing competing local access mechanisms, it will be fundamentally important to ensure, that competition in the local loop develops and, that network operators of local loops are not at the same time the only service providers over those networks.

Of particular interest is unbundling the local loop. Unbundling entails the separate provision of access to the switch and to the copper wire: this allows alternative operators to use only the copper wire of the incumbent, to invest in their own switching equipment and thus bypass the switching infrastructure of the incumbent.

Bundling can in itself constitute an abuse under Article 86 (EU Law), in addition, however, refusing to unbundle where such unbundling would allow competitors to invest in infrastructure which would upgrade the narrowband copper telecommunications network to broadband capability could, depending on the circumstances, constitute a separate abuse under Article 86(b) - that of limiting production, markets or technical development.

(v) Competition Cases

Within this new framework, competition policy is increasingly being applied to deal with antitrust and merger cases Already some major telecom competition cases came up in recent years On the basis of Article 85 (anticompetitive agreements), Article 86 (abuse of dominant positions, including issues of unfair pricing and refusing access and interconnection), and the Merger Regulation, were the basis for examining the planned mergers or alliances. Some of the cases were only acceptable from a competition point of view with sufficient remedies. Increasingly competition cases involve telecom companies and companies from neighboring sectors, indicating the emergence of convergence.

(vi) Telecom Framework

Competition policy will have an increasing role in telecoms markets and in markets with converging services. Sometimes the conclusion is therefore drawn that sector specific regulation must be gradually replaced by the application of competition law. Indeed, the application of competition rules can correct anticompetitive developments or interpret initially unforeseen challenges according to the spirit of the framework However, competition policy will not entirely replace regulation.

The issue is not one of specific sector regulation versus competition rules, but rather which evolution of the existing regulatory framework should be envisaged. This evolution will be accompanied by competition cases based on the experience they carry.

REFERENCES

Arthur, W. (1989). "Competing technologies, increasing returns, and lock-in by historical events," *Economic Journal*, *99*, 116-131.

Besen, S. M. & Farrell, J. (1994). Choosing How to Compete: Strategies and Tactics in Standardization," *Journal of Economic Perspectives*, *8*, 117-131.

Choi, J. P. (1994). Irreversible Choice of uncertain technologies with network externalities, *Rand Journal of Economics*, *25* (3), 382-401.

Choi, J. P. (1997). "Herd Behavior, the Penguin Effect, and the Suppression of Informational Diffusion: An Analysis of Informational Externalities and Payoff Interdependency," *Rand Journal of Economics*, *28*(3), 407-425.

Easley, D. & Kleinberg, J. (2010). Networks, Crowds and Markets, Cambridge: Cambridge Univ. Press.

Economides, N. (1996). "The Economics of Networks," *International Journal of Industrial Organization*, *14*, 673-699.

Farrell, J. & Saloner, G. (1992). "Converters, Compatability and Control of Interfaces," *Journal of Industrial Economics*, *15*(1), 9-35.

Gottinger, H. W. & Takashima, M. (2000). "Japanese Telecommunications and NTT Corporation: A Case in Deregulation," *International Journal of Management and Decision-Making*, *1*(1), 68-102.

Gottinger, H. W. (2013). “Chinese Innovation and Competition: An Industry Case of the Global Telecommunications Equipment Industry,” *International Business and Economics Research*, *2*(2), 22-32.

Gottinger, H. W., Umali, C. & Floether, F. (2010). eds., Strategic Alliances in Biotechnology and Pharmaceuticals, New York: Nova Science.

Grove, A. (1996). Only the Paranoid Survive, New York: Doubleday.

Hurwicz, L. (1999). “Revisiting Externalities,” *Journal of Public Economic Theory*, *1*(2), 225-245.

Katz, M. & Shapiro, C. (1986). “Technology Adoption in the Presence of Network Externalities,” *Journal of Political Economy*, *94*, 822-841.

Laffont, J. J. & Tirole, J. (2001). Competition in Telecommunications, Cambridge, Ma.:MIT Press.

Shapiro, C. & Varian, H. (1999). Information Rules – A Strategic Guide to the Network Economy, Boston: Harvard Business School Press.

Shy, O. (2001). The Economics of Network Industries, Cambridge:: Cambridge Univ. Press.

The Economist (2000). “Untangling E-Economics: A Survey of the New Economy,” Sept.23, 5-47.

Weitzel, T. (2004). Economics of Standards in Information Networks, Heidelberg: Physica Verlag.

Chapter 3

DYNAMICS OF COMPETITION AND MARKET DYNAMICS

Largely from the perspective of network economics we look into a broad spectrum of economic studies on competition, innovation and technological change developed primarily over the last three decades with a view to relating theories, models and ideas on how best to analyze the impact of innovation and market dynamics. Innovation has a major network inducing effect, so has competition. Competition in this context always carries dynamic features and is intrinsically linked to games of strategy in the original von Neumann-Morgenstern version (Case, 2008). Our review focuses on a range of studies on technological change and discusses both prevailing models of innovation and studies that fall within other traditions such as evolutionary economics.

Here we start with a discussion of primary concepts from which the development of assessment is taking shape. In the view of McAfee (2002) they belong to the 'strategist's toolkit.' They are largely drawn from an overlapping range of the industrial organization and endogeneous growth literature.

3.1. MARKET DYNAMICS AND COMPETITION

We offer a structured approach to the analysis of the impact of technological change on competition and to the consideration of innovation in antitrust or merger and acquisition (M&A) cases. While these draw on the ideas from analytical and theoretical work, they seek to provide strategic considerations of application value given the inevitable constraints of competition policy analysis (Motta, 2004).

In particular, we reach the view that explicit consideration of market dynamics may be particularly important in the context of a competition policy analysis for two reasons: Dynamic markets tend to be 'unstable.'

Competition in innovation features prominently in the competitive process. The standard legal practice of competition policy analysis places emphasis on firms' positions and market shares to assess competition within current relevant markets. This is not sufficient to capture the effect on the competitive process as a whole. We need to consider both the effects on product market competition and the effects on competition in innovation.

3.2. Competition Analysis

Dominance and abuse of dominance in dynamic markets may often be assessed if the conduct in question impedes competition by restricting other firms' capabilities to innovate and supply future markets – as some cases for diverse network industries appear to show (Kwoka and White, 1999; Spulber and Yoo, 2009; Shelanski, H. A. and J. G. Sidak, 2000).

In particular, a firm that holds a dominant position in the supply of an 'access service' on which other firms depend if they are to innovate may be in a position to distort competition in innovation. Similarly, a dominant firm may take other actions that are shown by the analysis to have inhibited other firms' capabilities to supply future markets.

In such cases we emphasize how the market definition, their scale and scope and capability to supply future markets can be used to assess the effects of the merger or agreement on (a) product market competition and (b) competition in innovation. As regards product market competition, a forward-looking perspective may need to consider the competitive effects in the future 'evolution' of the current relevant markets. It may also need to consider the expected competitive effects in separate future markets relevant to prospective products which are yet to be introduced.

As regards competition in innovation, a forward-looking assessment may need to consider not just whether competition will be impeded in the supply of current and anticipated products, but also whether competition will be impeded in innovative activity associated with potential products that are yet to be developed.

3.3. Sources of Innovation and Competition

It has been only recently, mainly over the last four decades, that the economic forces behind technological innovation have started to be investigated in more detail within modern economics. As a result, a wide variety of theories and models, sometimes very diverse in spirit, describing the economics of innovation are now available. All these theories share the common aim of providing a conceptual foundation for understanding how innovation affects the economy, how economic forces affect the emergence of technological changes, and the decision-making processes through which technological innovation occurs.

Traditional models of innovation focus on the study of firms' incentives to invest resources in Research and Development (R&D) activities. Game-theoretical models (GTMs) developed in industrial organization have investigated firms' R&D decisions in strategic environments (Gottinger, 2006).

GTMs in networks (Easley and Kleinfeld, 2010) – being genuinely competitive – suggest that there are two main forces that underlie firms' investment in R&D: the search for higher profits and the threat posed by falling behind potential innovating rivals. They study these forces in a variety of market situations and address issues such as the interplay between innovation and market structure, the dynamics of competition and the nature of the relationship between intensity of competition and innovation.

These models provide a rich picture of what the plausible strategies and industry equilibria in dynamic markets can be. However, general predictions, that can be considered appropriate across all situations and industries, are scarce. On one hand, this is a result that seems to stress the lack of predictive power of these models, i.e., (almost) everything can be rationalized; on the other hand, however, the variety of results seems well to fit with the variety of observed behaviors and 'equilibria.' There is no general model that can uncritically be applied to any case: they have to be adapted to specific market situations.

Despite the absence of general results, these models are certainly useful tools to understand firms' incentives to invest in R&D activities in strategic environments and to suggest what main factors may be central in shaping the nature of dynamic competition. For instance, these models suggest that: in order to understand R&D investments it is necessary to understand how innovation may affect profits both of successful and unsuccessful innovators. One perspective captures the idea that firms want to innovate to increase their profits, another that firms want to innovate to maintain competitiveness.

The relationship between concentration of an industry and its rate of technological innovation is certainly complex and in general not a causal one: both should be thought of as the outcomes of the operation of market forces and exogenous factors such as the nature of demand, technological opportunity and specific industry related conditions.

Dynamic competition may be characterized by persistent dominance of the incumbent leader or by action-reaction whereby incumbents are overtaken by a rival whose incumbency itself could then be short-lived. The nature of market dynamics depends on a number of factors, such as the type of innovation, i.e., drastic or non-drastic, the uncertainties involved in R&D activities, the nature of patent protection and of knowledge spillovers, the intensity of product market competition, among others.

When the relationship between competition and innovation is investigated, it is necessary to be clear what the notion of intensity of competition describes and how this relates to market structure. Indeed a market where competition is tougher may be more concentrated simply because inefficient firms cannot survive. There may be trade-off between the intensity of static competition and innovation. In general, the relationship between intensity of competition and innovation need not be monotonic at all.

Endogeneous growth models suggest that innovation, resulting from intentional R&D investments by profit-seeking firms or simply by unintentional learning-by-doing, is a fundamental driver of economic growth in the long run.

Early endogenous growth models stressed the importance of ex-post rents for innovation: competition would have a detrimental effect on innovation by decreasing the rents that an innovator would be able to gain. More recent models have emphasized another mechanism by which competition affects innovation: tougher competition may increase the incentives of firms to innovate in order to escape from fierce competition (see Tirole, 1988, Chapters 8-10; Aghion and Griffith, 2005). More recent studies suggest that the relationship between competition and innovation may not well be monotonic and that instead, one should expect an inverse-shaped relationship: when competition is low, an increase in competition would foster innovation; the reverse would happen when competition is fierce.

The result that competition may be conducive to innovation is also obtained in studies where the traditional behavioral assumption of profit-maximizing firms is relaxed. When principal-agent considerations are introduced to explain managers' behaviors, another mechanism by which competition may favor innovation is suggested: the speed of innovation may be retarded by the slack of managers who tend to avoid private costs associated with innovation. When competition intensifies, the higher threat of bankruptcy may force managers to speed up the process at which new ideas are adopted. Hence, competition may be conducive to faster rates of innovation.

The evolutionary approach to the study of innovation has been developed on very different methodological basis than those underlying traditional economic models of innovation. In particular, we observe the rejection of the modeling assumptions of rationality and equilibrium that are fundamental to the traditional approach.

Evolutionary economics looks from the outset at dynamic processes. In particular, it is associated with the use of analogies from evolutionary biology to explain economic growth and the process of competition. Thus the cornerstones of an evolutionary analysis of competition and innovation are variety, selection and imitation (Jacquemin, 1987, Chapters 2, 3). At a basic level, using Darwinian analogies, we can begin to appreciate the role of the market in selecting the more fit firms (efficient and profitable), products and techniques at the expense of less fit firms. In addition to this effect, we would expect to see imitation of winning ideas by those whose survival is otherwise threatened (although this is limited by the tacit nature of knowledge).

Inherent in this model of competition is the association between competition, experimentation and variety. A variety of experiments allows, through the process of selection, far greater economic progress than would be available through uniform optimization.

The economic development and innovation can be seen as a combined effect of selection (via competition) from a variety of competing routines and practices as well as the more endogenous process of agents seeking improved routines and practices. While the latter is certainly incorporated, if treated somewhat differently, under the mainstream neoclassical approach, the emphasis on selection from variety seems an important addition to this approach.

This suggests that models found in other branches of economics might miss something important when they analyze dynamics with reference to homogenous profit-maximizing firms: namely the benefits of selection from heterogeneity in capabilities and innovative experimentation.

3.4. Competition for Network Industries

Product market competition can be understood as competition between firms in the supply of existing products. This competition can be conceptualized as the rivalry between firms in terms of marketing, and notably pricing, of their products taking as given the characteristics of products (including production processes and costs). In markets that are considered relatively ‘static,’ product market competition is the main channel through which the competitive process takes place.

Competition in innovation or ‘innovative competition’ (Fisher et al.,1983, 38) can be understood as the competition between firms to develop new products and production processes; this competition is often associated with the ideas of a competitive threat from innovation. Two firms are competing in innovation if they are undertaking (uncoordinated) innovative activity that can be identified with the prospect of introduction of products or services that will compete in the future. This innovative activity could be investment in R&D

or less formalized activity such as product improvement through a process of learning-by-doing. Where two firms are competing in innovation, we expect their decisions regarding innovative strategy to be influenced significantly by the innovative strategies employed by rival firms. Failure to innovate successfully will lead not just to loss of potential profit but also risk falling behind innovative rivals (Brenner, 1987, Chapter 3).

Traditional industrial economics. as covered in essays by George Stigler(1968), also makes use of the terminology of 'static' and 'dynamic' competition. Along the same lines as above, the static dimension of competition is usually related to price competition, which takes place given the set of products or services that are marketed by firms. Dynamic competition, on the other hand, refers to the process whereby new and improved products, services and production processes are introduced. However, in the context of 'innovation and competition' we consider it more appropriate to use the terminology of 'product market competition' and 'competition in innovation' as this is more specific and more conducive to the development of applied competition analysis and policy.

Winner-Takes-All Markets

Industrial organization models, and the review of studies of network industries, suggest that competitive interaction in innovative activities may take different forms according to the structure of payoffs of innovative activities to 'winners' and 'losers.'

In some organization models, and the review of studies of network industries, suggest that competitive interaction in innovative activities may take different forms according to the structure of payoffs of innovative activities to 'winners' and 'losers.'

In the first case, competition in innovation may be the essential, or at the limit the only dimension of the competitive process, i.e., winner-takes-all markets. Persistence of monopoly may be observed in a winner-takes-all market, but provided that competition in innovation is effective, this does not necessarily imply that competitive forces are muted. On the other hand, in these markets, competition in innovation is the area where the current dominant firm may be more likely to abuse its dominance (since, by definition, such markets are not conducive to sustained product market competition).

These considerations suggest that it is useful to distinguish between different economic environments, according to the relative importance of competition in innovation and competition at the product market level that can be expected in the market. The literature on network industries is particularly useful in explaining why a market may exhibit winner-takes-all properties.

Network Effects

Network effects may be an important factor in determining the competitive environment.

A (positive) direct network externality exists when the demand for a service increases by an extra unit. In order for this effect to be 'direct' the reason for the increase in demand must come directly from the additional consumption of the service in question, without need for a strategic response by suppliers in the same or related market. For example, the demand for public telephony services (i.e., the demand for subscriptions to networks) may increase as more users consume this product (e.g., enter into subscription agreements with public network operators) simply because more users allow a greater number of potential connections to be made of the network.

By contrast, a (positive) indirect network externality exists where the demand for a service supplied in one relevant market has a significant effect on supply of another product (typically in a different relevant market) such that this in turn increases demand in the first relevant market.

An example of indirect network effects is that of computer software for a particular operating system platform. An increase in consumption in the market for supply of operating systems is likely to bring benefits to the markets for supply of compatible software, by expanding the market such that a greater variety of software is offered or the unit price decreases (since fixed costs of software development can be spread over a greater number of potential buyers). In turn, the benefits to the software markets render the same operating system more valuable, and thus increase demand for it. But the effect is indirect because it relies on a strategic response of software suppliers to the increase in the consumption of operating systems.

Horizontal and Vertical Innovation

The study of industrial organization and endogenous growth models suggests that the distinction between ‘horizontal’ and ‘vertical’ innovation may be useful. Horizontal innovation entails the discovery of a new product which, setting aside price considerations, is considered better than existing products only by some users (or for some uses). By contrast, vertical innovation entails the discovery of a new product which, setting aside price considerations, is considered better than existing products by all users (for all uses); hence the idea that products can be ranked according to a ‘quality ladder’ (Grossman and Helpman, 1993, Chapter 4).

The importance of this distinction derives from the differences in market dynamics that are associated with the two types of innovation. In particular, horizontal innovation generally results in the creation of new product groups that can coexist with older product groups.

Vertical innovation, on the other hand, is generally associated with a process whereby new and better products displace older obsolete products from the market. In markets subject to vertical differentiation, a firm (or more accurately a product) that does not follow the pace of technological advance may be driven out of the market. In other words, in order to survive, a firm needs some basic capabilities, and the need to improve these capabilities over time as its competitors improve theirs (hence the idea that a rising ‘quality window’ exists, outside which firms cannot survive in the market). This economic process can guide on the nature of market dynamics that may affect a particular industry.

Step-Wise Innovation versus Incremental Innovation

Step-wise innovation involves a relatively substantial degree of novelty, e.g., a new product or production process that is substantially different and/or better than older products.

By contrast, incremental innovation is characterized by minor cumulative changes to products or production processes.

The extent to which an innovation is novel may be an important factor in assessing whether incumbent firms have an advantage over potential entrants or vice versa. In fact, industrial organization models suggest that incumbent firms may have different incentives to innovate than entrants. Similarly, given the difficulties that incumbent leading firms may have to deal with drastic changes as a form of of intrafirm systemic inertia, the opportunities for step-wise innovation may relatively favor incumbent firms.

Intensity of Competition and Innovation

There is little consensus about the relationship between intensity of competition (however defined) and innovation. At a general level, there is some evidence in favor of an inverse shape, indicating that innovation may be relatively less rapid at both very low and very high intensities of competition. However, there is the question of at what point is more competition worse? This cannot be answered at the generic level, given the current level of understanding of these effects. Many of these problems derive from the difficulties involved in understanding the link between product market competition and innovation. This issue is best addressed on a case-by-case basis.

Firms' Heterogeneity and Capabilities

Prevailing models of dynamic competition deal essentially with firms' incentives to innovate. Not much attention is paid by theories rooted in the neoclassical tradition to the resources that firms need to innovate, nor to the importance that firms' heterogeneity may have in affecting market dynamics. For instance, whether a market is prone to dynamic persistence of monopoly depends not only on the incentives to innovate that an incumbent firm has compared to a potential entrant, but also on their relative capabilities, which may not be immediately ostensible.

The notion of firms' capabilities seems crucial to the understanding of market dynamics driven by innovative activities. In fact, we believe that it is by emphasizing the importance and nature of different firms' capabilities that evolutionary theories of innovation are most useful to apply.

Capabilities to innovate may depend on a number of factors: having (the ability) to access external sources of knowledge (e.g., scientific knowledge); having access to some tacit knowledge which is not transferable (perhaps because embodied in organizational routines); and having access to some specific complementary resources, e.g., specific input factors.

Tacit Knowledge

The capability of a firm to innovate in particular areas will be affected by the extent of tacit knowledge supporting its innovative activity, and whether the firm has access to this knowledge (e.g., acquired through a process of learning-by-doing).

Tacit knowledge cannot be codified, transferred or imitated by external parties but is rather embodied in routines and skills of the firm. The term 'tacit knowledge' is used to encapsulate the idea of knowledge that is found to be useful (e.g., in an organizational routine or a skill) without being articulate, and therefore without being directly transferable (e.g., sold as information).

3.5. Dynamic Markets and Competition

Every market is dynamic to the extent that we expect changes over time, for example in the quality and characteristics of the goods and services produced, their prices, the processes used in their production and the firms engaged in their supply. Innovative actions of firms affect all markets in the economy.

In line with the fundamental distinction we have drawn between product market competition and competition in innovation, we discuss how a more explicit consideration of market dynamics can inform on each of these dimensions of analysis.

First on the product market dimension, in dynamic markets time matters. A snapshot of the market observed at the time of the investigation is likely to be very different from a hypothetical snapshot of the same market taken in the near future. Such changes may be due to innovation undertaken by the firms in the market, or due to factors outside the market such as external technological developments, or shifts in consumer tastes. The result may be either more competitive or less competitive market conditions over time.

Second, in dynamic markets, the production, assimilation, and commercial use of new knowledge is central to the competitive process, i.e., it is a fundamental determinant of firms' success and failure. Thus, in some cases innovation (i.e., activity related to the supply of new services in the future) may be a crucial dimension along which the process of competition takes place.

Of course, anticipation is subject to great uncertainties. The view that the markets of telecoms, computers etc. converge was largely anticipated but it turns out to be different than expected and therefore may run contrary to anticipated market definitions. Take an example in telecommunications: G3 was the name of the game about two decades ago where billions of euro (equivalents) through UMTS were skimmed off from anticipating market participants, only to know later that rival standards based on new technologies (Wi-Fi) bypassing UMTS, would create completely new markets not being anticipated.

Market Dynamics and Product Market Competition

If analysis of competition in a current relevant market is to provide guidance on dynamic effects in that market, we require some understanding of how this current relevant market is likely to evolve in the future. In areas of economic activity where the relevant markets are expected to remain stable, analysis of competition in the current relevant markets, with due attention to likely entry to that market, could be broadly informative on future competition in those markets.

However, market dynamics will often mean that analysis of competition in the current market may not be a reliable indicator of competition in the future. Technical change may affect market definition, for example by widening the range of possible substitutes. In addition, technical change may affect the structure of the market as defined, for example, by making economies of scale more or less important. Such changes may be due to innovation undertaken by the firms in the market, or due to factors outside the market such as external technological developments, or shifts in consumer tastes.

Taking account of these factors is not simply a question of examining entry possibilities. Indeed it requires consideration of how the market to which entry applies, and the barriers to entry that may exist in this market, changes over time.

This concern has various practical implications. For example, in analyzing a merger, the two merging firms may currently be operating in completely separate relevant markets, thus raising no current competition concern. But product development may mean that the relevant markets converge in the future, such that these firms are expected to be competitors in the future. This does not mean the merger should be blocked; simply that competition analysis may need to be forward-looking as to how current markets are likely to evolve in future.

Furthermore, in an extreme case of the ‘stability’ problem highlighted above, a firm may not be active on any current markets but may be developing products that would compete in the future.

This may be the case for a merger between biotechnology companies that have no current products to sell – as shown for the history of this industry (Gottinger, Umali and Floether, 2010). In this case the competition examination may need to explicitly define the future markets on which the outcome of the parties’ product development would be traded, in order to analyze the effects of the merger on product market competition. The aim would be to establish whether the expected outcomes of the biotechnology companies’ product development programs - i.e., the products the parties are anticipated to supply in the future - would face product market competition from other suppliers when they are introduced. This analysis would abstract from the effects of the merger on the parties’ development of these new products; thus as an additional stage the analysis of this merger would need to consider the effect on competition in innovation, which is discussed below.

Competition in Innovation

By focusing on competition in defined product markets, be these current or future, analysis may miss the effect of competition in innovation. In some instances this will not hinder successful examination of a competition case, either because competition in innovation is subordinate to competition in the product market, or because analysis of the effects on competition in the product market may act as a ‘proxy’ for analysis of the effects on competition in innovation.

However, where innovation is clearly an important part of the competitive process, the effect of a merger or anticompetitive conduct on competition in innovation may be significantly different to its effect on competition on the product market. For example, if a market seems to exhibit ‘winner-takes-all’ properties, such that it can only sustain one firm at a point in time, an abuse of dominance case may need to explicitly consider whether conduct by the dominant firm affects competition in innovation rather than product market competition.

As such, competition in innovation may need to be considered in its own right.

3.6. Current and Future Markets

Market definition beyond scale and scope of innovation activity (Lerner, 2012) is also an important element of merger and anti-trust investigation and seeks to identify the competitive constraints that derive from consumers' substitution patterns (Whinston, 2006,77). Innovation that affects products' characteristics and prices is a major source of instability of consumers' substitution patterns, and hence of the boundaries of relevant markets over time.

Furthermore, market definition, by making use of the concept of a 'future market,' can allow for conceptually robust analysis of the effects of M&As on both prospective product market competition and on competition in innovation.

Competition Law through Article 82 of the European Union is supported by the view that the unilateral behavior of a firm can only significantly distort the process of competition to the detriment of consumers where that firm holds market power (Nouel, 2001, Title II).

It is necessary to consider where a firm holds market power in order to establish where it can distort competition (in the same or a related market). In the context of an agreement or merger, market definition allows for identification of the competitive constraints that might be lost if the merger were allowed to proceed.

Competitive Issues that Arise in Dynamic Markets

The purpose of market definition is to identify a relevant market as those products and services, the suppliers (or potential suppliers) of which are capable of exerting effective competitive pressures on each other and of constraining each other's behavior.

The major challenge that market dynamics pose to market definition derives from the instability of the market environment: technological change alters the set of products or services sold, how they are produced, their characteristics and prices and hence affects substitution patterns and the related competitive constraints. This instability may be most profound when the markets considered are reflecting a substantial change in the underlying technological paradigm but can also derive from less drastic change along a given technological trajectory.

This process will have implications for the appropriate assessment of the effects of an M&A by a dominant firm (not to speak of a 'collusion') on competition between products currently on the market (Marshall and Marx, 2012).

Changing consumers' substitution patterns imply that the temporal aspect of market definition is likely to be particularly important. Most clearly the boundaries of the current relevant market will evolve, perhaps due to changes in product characteristics and relative prices. Consequently, analyzing whether two products are substitutes today may provide a poor guide of whether they will be substitutable, and hence likely to compete, in the future. Furthermore, expectation that new products will be introduced in the future may provide a constraint on the terms under which current products are supplied. This suggests the potential importance of using the concepts of a future market to consider how the current relevant market (i.e., the traditional starting point for competition analysis) can be expected to evolve over time.

In addition, competition assessment of an M&A by a dominant firm may need to pay attention to the competitive effects on products not yet supplied to the market, but whose introduction on the market can be anticipated to some reasonable extent; this has implications for market definition analysis. For example, in analyzing a merger in a dynamic setting, it may be necessary to consider not only the future evolution of markets for the merging firms' current products, but also the relevant future markets on to which the firms are expected to introduce new products.

Demand substitutability is also affected by changes in products' attributes and by changes in the available set of products. Dynamic competition that results in such changes may hence affect the boundaries of competition between firms and the boundaries of relevant markets. For instance, a new technology may, due to its insufficient quality, initially be considered not to be a substitute for the old technology but technological improvements may in the near future cause the two technologies to be considered close substitutes by consumers. Conversely, substitutability between a new high-quality/high-price product and an old low-quality/low-price product may increase over time as relative prices converge sufficiently to reflect consumers' perceptions of quality differences.

In some cases, technological change is expected to result in new relevant product markets altogether. For example, this may be appropriate in analysis of a merger between two pharmaceutical firms that are investing in R&D targeted at the discovery of new drugs for the same currently untreatable disease. In this case the new drugs would not be expected to face competition from existing drugs and would probably constitute an entirely new relevant market. If estimates can be made of future product attributes then some notional 'future market' can be considered by applying the principles of the hypothetical monopolist test not to the current supply of products but to the expected future supply of products. As far as assessment of the effects of a merger, agreement or conduct by a dominant firm is concerned, there are two general reasons why it may be appropriate to consider future markets:

(i) to allow forward-looking assessment of the competitive effects on the supply of current products in the future, in light of anticipated or potential market developments (this can be considered as the 'future evolution' of the current relevant markets),
(ii) allow assessment of the competitive effects on the supply of products that are yet to be introduced, but whose introduction can be anticipated and is relevant to the investigation.

Therefore, future markets can be defined both from the perspective of the supply of current products (looking forward) and also from the perspective of the supply of products whose introduction is anticipated. The latter may be relevant, for example, in analyzing a merger between two firms that have potential products in the 'R&D pipeline' that are yet to be introduced but which can be identified with some confidence; the potential products expected to result from the pipeline could be associated with the loss of expected competitive constraints in future product market competition.

The time horizon to be chosen for this future market would depend on the specific characteristics of the case, e.g., the pace of technological progress. While we cannot capture all the possible future markets within the practical constraints of investigation — in terms of timing and information— there is benefit in attempting to consider those future markets where we expect the effect of a merger (agreement, or conduct) to be significantly different than in the current relevant market.

However, there is a caveat that emerges from analytical considerations on future markets as expressed by Arrow (1974, 6) in his Presidential address to the American Economic Association:

> "The implication first of all is that the information needed by the optimizer is not provided by an existing market. It will be provided by a market which will exist in the future, but that is a bit too late to help in decisions made today. Hence the optimizer must replace the market commitment to buy or sell at given terms by expectations: expectations of prices and expectations of quantities to be bought or sold. But he cannot know the future. Hence, unless he deludes himself, he must know that both sets of expectations may be wrong. In short, the absence of the market implies that the optimizer faces a world of uncertainty."

New and Old Technologies

Assume a launch of a new technology: a niche of a new relevant market for current supply of new technology. Now consider new technology diffused: a single relevant market for new and old technology in which a relative price difference is related to relative quality difference (if the two technologies coexist).

The potential for such a scenario may be of particular interest in considering abuse of dominance cases. A firm supplying the old generation product may appear dominant today because consumers are not willing or able to substitute between the old and new products today. However, this supplier may not be dominant at all as it could be strongly constrained by the threat that its customers will switch to future consumption of the new technology, once it has diffused and its price has fallen. Such a case would demonstrate both the value of considering the time dimension of the market and the complexities of assessing dominance in dynamic markets.

Inter-Temporal Substitutability

Analysis of the effect of innovation on the scope of the relevant market may need to define both the current market(s) and also future market(s).

In defining the current relevant market, analysis would start by consideration of whether the price of the old technology is constrained by the price of the new technology at the current point in time, taking their relative qualities and prices as given. This assessment may find that the new technology is introduced at a very high price such that consumers of the old products would not readily substitute even though there is a quality advantage. This would suggest new and old technologies would be in separate markets.

However, we also need to consider the time dimension of the market. While current supply of the new technology may not constrain supply of the old technology, consumers will take into account their expectations of future prices. In this example, they expect a substantial fall in the price of the new technology. This expectation may mean that they are prepared to substitute between buying the old technology today and waiting to buy the new technology in the future. Thus a hypothetical monopolist of the old technology (today) would be constrained in its pricing by the threat of consumers delaying consumption with the specific intention of buying the new technology in the future, even if it is not constrained by today's supply of the new technology. This is a genuine constraint on the market power of the hypothetical monopolist. Therefore the relevant market should be defined to include the current supply of the old technology and the future supply of the new technology.

This relevant market may then be accompanied by a 'niche' relevant market for those buyers of the new technology who value this far higher than the current technology and are not prepared to wait for the future technology to drop in price, and therefore do not substitute over time. These consumers would comprise the bulk of purchasers at the initial release of the product, in what would be a separate relevant market to that defined above.

Nonetheless, in defining today the future market (i.e., one outside the time window of the current relevant market), based on the expected prices in the table above, we could well expect there to be a single relevant market for the supply, at that future point in time, of the new and old technology. This would be the case where the relative price difference between the technologies has fallen to a level that reflects the perception of the quality difference between the two products, in which case they would be generally regarded as substitutes.

3.7. PRODUCT SUBSTITUTABILITY

How will technological progress affect the boundaries of the relevant market over time by considering the impact that expected changes in products' characteristics have on consumers' substitution patterns?

The analysis could proceed at two levels: the identification of the different technologies associated with the products competing in the market; and the analysis of the directions of innovation within each technology ('technological trajectory').

For instance, if cameras were considered, the first step of the analysis would identify the existence of two (main) competing technologies, a conventional technology and the new digital one, and the second step would consider the trajectory of technological change along which digital cameras are subject to innovation, e.g., increase in the resolution of pictures. The likely evolution of market boundaries may then be analyzed on the basis of the expectation of a process of diffusion of the digital technology connected to the process whereby a better (and improving) technology replaces an inferior one, although elements of horizontal differentiation between the two technologies would suggest that cameras embodying different technologies are likely to coexist in the market despite this diffusion process.

The analysis of technological change and product substitutability should result in the understanding of the likely evolution of market boundaries needed to define current and future markets as sketched along the following line:

Identification of current products $\Rightarrow$ Analysis of technologies/technological trajectories $\Rightarrow$ Choice of time horizon $\Rightarrow$ Definition of future relevant market.

The ability of assessing, with any degree of confidence, the evolution of market boundaries relies crucially on the capacity of identifying past techno-economic trends that are expected to continue in the future. This would be less likely if a radically new technology has just been introduced in the market but would be more feasible when there is evidence of ongoing directions of changes such as when innovation is of a more incremental nature.

The specific characteristics of each case would clearly determine the time-horizon of the analysis, which also crucially depends on the extent to which confident predictions on the evolution of technological change can be reached.

In any case, it is important that the analysis is undertaken with focus on the consumers' perspective, and that a broad approach is followed since even changes that appear external to the relevant market may be important in affecting the scope of the market in the future.

In order for products' changes to be considered indicative of the nature of future demand substitutability, the consumers' perspective is adopted: no change of market boundaries may be expected to take place if there is no likelihood of changes in consumers' patterns of substitution between different technologies/products.

Moreover, it should be noted that technological changes taking place outside the current relevant market might also have an important impact on its evolution, i.e., on the identification of future relevant product markets. For instance, technological change that affects the products in a different market than the relevant one may impact on the latter due to convergence of products' characteristics and functionalities (e.g., mobile phones, smartphones and tablets). Hence, it is necessary that a broad approach is followed, which considers intra-market but also broader technological trends.

Technological Trajectories

Market dynamics may result not only from the introduction and diffusion of products that embody a new technology in the market but also from technological change that takes place within a given technology, along what we call a 'technological trajectory.'

If such technological change were to proceed randomly, along directions that cannot be predicted with any degree of accuracy, it would not be possible to make any reasoned statement on the evolution of the boundaries of product markets over time. Definition of future markets would be subject to such uncertainty as to render it no use to competition assessment.

However, it is often possible to observe a relatively ordered techno-economic pattern of innovation at the level of single technologies. For instance, Dosi (1988) observes that technological progress in aircraft technology

> "…has followed two quite precise trajectories (one civilian and one military) characterized by log-linear improvements in the trade-offs between horsepower, gross takeoff weight, cruise speed, wing loading, and cruise range."

The notion technological trajectory is a useful one since it suggests that it is possible to identify the major trends that characterize the evolution of a technology in a certain market and to predict their impact on market boundaries.

Within each technology, it is likely that a technological trajectory is identifiable in terms of the pattern of changes that characterizes innovation within that technology. The analysis of technological trajectories may allow the identification of whether or not innovation is resulting in convergence between currently separate markets or, vice versa, in fragmentation of a current relevant market into a number of different future markets.

It is very hard to provide guidance at this very general level, since the evolution of dynamic industries rarely follows predictable patterns. Even the life-cycle model that has often been considered a valuable description of a representative pattern seems not to fit all industries. Nonetheless, by considering the nature of the innovation we can form some view on the likely effect on the scope of the market, and hence competition. For practical reasons, the analysis may proceed by identifying a set of key characteristics that are likely to be key drivers of the evolution of substitution patterns (e.g., the improvement of the resolution of digital cameras) and are central to the technological trajectory in the market. Quantitative evidence on the past trends of the evolution of these key performance or design variables may be useful to infer the likely evolution, included the time dimension, of changes of products' characteristics.

3.8. Issues with Market Definition

Market definition seems, from an economic perspective, to be a necessary component for establishing dominance in Article 82 cases of EU Competition Law, be they static or dynamic. Article 82 is supported by the view that the unilateral behavior of a firm can only significantly distort the process of competition to the detriment of consumers where that firm holds market power. The moment we recognize that each firm undertakes different activities, supplies different products, competes in multiple dimensions, and evolves over time, the importance of market definition becomes clear. The extent or broadness of market definition indicates how firms could position themselves as market leaders or 'champions' (Simon, 2009, Chapter 3). Since firms may be dominant in some markets and not dominant in others, market definition sets out the areas of economic activity where a firm may be able to commit abuse. It is necessary to consider where a firm holds market power in order to establish where it can distort competition. The distortion to competition could be either in the market in which dominance is held, or in supply or innovation associated with a related market. On balance, misuse may only characterize negative aspects of market power and need to be determined on a case-by-case basis. For example, gaining market power through innovation is a positive incentive and thus by itself not problematic and need to be targeted at repeated misuse to suppress competition but only in cases which are not based on merit or technological preeminence (Gottinger, 2003, Chapters. 5, 6).

In the context of a merger, market definition allows for identification of the competitive constraints that might be lost if the merger were allowed to proceed. This is important in determining the extent to which the merger is likely to impede competition, taking account of any pro-competitive gains. Indeed market definition is especially valuable in dynamic markets where the competitive conditions under which firms operate vary across the range of services they supply (or could supply), vary across the supply chain in which they are active within and vary over time. Only through robust market definition, and understanding of the relationships between different markets, can the implications of this variation be clearly set out.

Moreover, as highlighted in the introduction, for two firms to be placing a competitive constraint on each others' innovative activity (i.e., competing in innovation) implies that we expect, with some non-negligible probability, that they will introduce products to the same future market. Thus market definition in the context of innovation provides an analytical framework through which we can determine, for example, whether the innovation undertaken by two firms can be seen as competitive. If we do not anticipate that the outcomes of the innovative activity — be these final products or some intermediate goods — will be substitutable in the eyes of customers, then it is difficult to argue that the firms are competing in innovation. The thought experiment behind the hypothetical monopolist test provides a consistent approach for this assessment, regardless of whether data exists for a quantitative 'test.'

Finally, some argue that a weakness of the hypothetical monopolist test in dynamic markets is that it leads to the definition of markets that are 'too narrow.'

Concerns with findings of narrow markets may be valid where definition of a narrow market is combined with a view that dominance per se is bad (when, in fact, anticipations of gaining market power often provide a necessary spur to innovation) or when 'narrow' markets are used to inappropriately infer dominance (e.g., if too much emphasis is placed on market shares and profit analysis and too little on the scope for entry or expansion by other firms). Ultimately such concerns, if well founded, call for a better assessment of competition in the market. But this is not a reason to change the concept of a relevant market for dynamic cases.

Innovation Markets

The IP Guidelines in the US (1995) introduced the concept of an 'innovation market' as an analytical tool to consider the competitive effects on innovation and R&D, rather than identified future product markets:

> "An innovation market consists of the research and development directed to particular new or improved goods or processes, and the close substitutes for that research and development. The close substitutes are research and development efforts, technologies, and goods that significantly constrain the exercise of market power with respect to the relevant research and development, for example by limiting the ability and incentive of a hypothetical monopolist to retard the pace of research and development. The Agencies will delineate an innovation market only when the capabilities to engage in the relevant research and development can be associated with specialized assets or characteristics of specific firms…."

The innovation market concept relates to the research and development effort that is associated with the future introduction of innovations.

Insofar as any firm that is currently engaging in R&D and other innovative activity may have some tangible output that it could sell today, there are grounds for the concept of a relevant market for the supply of this 'innovation.' Where innovative activity has led to the creation of information, i.e., knowledge that can be codified (as distinguished from tacit knowledge), or even prototype products and designs, we could consider whether the firm holding this information would face competition if it offered it to market, supported by IP protection.

Indeed the case exists in practice. For example, in the pharmaceuticals sector, small biotechnology companies exist that supply intermediate products in the supply chain, such as molecules that are yet to be tested in clinical trials but have chemical potential to treat certain indications. Where two such firms seek to merge, competition policy analysis would benefit from consideration of the relevant market for this output. But this would probably be standard practice, since the conventional output of these firms is essentially an R&D outcome that is then used as an input by large pharmaceutical firms (see Gottinger et. al., 2010). There is no conceptual difference to this market for the supply of an intermediate input than a market for the supply of traditional manufacturing component inputs.

But even if a firm is currently undertaking R&D with no intention of disclosing or selling the information that is associated with each subsequent stage of product development - until it patents and introduces the final product that this R&D has led to - we can consider what competitive constraints it would face were it to offer this information to the market. For example, consider a vertically integrated firm that controls its supply chain from initial product development to final supply on the market. Even if this firm does not currently license any of the intellectual capital associated with the products under development, we could still ask what the relevant market might be were it to supply such a service, and consider whether there would be market power in this market for the potential supply of a license. This in turn would inform on the competitive constraints that the firm would face in the supply of this information. Such an approach might be necessary in abuse of dominance cases, where failure to supply some kind of access service denies competitors access to a related product market. The importance of identifying such 'access services' in considering the relationship between competition and innovation is discussed next.

However, such markets for the potential supply of information relating to the intermediate output of innovation are not the same as the 'innovation market' approach defined above. Both the examples above concern cases where the purpose of analysis is to analyze competition, or potential competition in the supply of a service. Such analysis has only a superficial difference to standard market definition and competition analysis, namely that the service under investigation is information related to a potential product, rather than a more tangible intermediate good.

In contrast, the US innovation market seems more associated with R&D input for its own sake, rather than in relation to some information output that an innovating firm could potentially trade. Moreover, the US definition goes beyond the R&D output that firms may sell on the market, to the identification of a concept of market power in innovation: any concept of market power must relate to a relevant market over which this power is held, even if this is defined implicitly.

Consider again the case of the biotechnology company selling IP-protected information. If this firm has market power it is because its price in supplying this information to potential customers would not be constrained by the supply of similar information, or the potential supply of similar services, by competitors. This market power can be understood in the normal way (expected lack of competitive constraints on pricing) without reference to whether the firm can 'retard the pace of research and development.' As argued elsewhere in Gottinger (2003), for practical purposes it seems hazardous to define market power with reference to whether a firm can retard the pace of innovation, simply because it may be that competition itself is the retarding force on innovation.

By creating a concept of market power (and therefore a market) that exists in relation to innovative activity (i.e., input), rather than in relation to the supply of intermediate or final products derived from that innovation (i.e., output), we risk confusion between competition in innovation and the effect of competition on innovation, and between competition in innovation and competition in the supply of the output of that innovation. Since competition in innovative activity is not itself associated with market power it does not seem appropriate to define a market for it.

Instead, we consider that the fundamental separation we draw between analysis of the effects of a merger, agreement or conduct on product market competition and the effects on competition in innovation is the most useful approach to follow. One of the key drawbacks of the innovation market approach is that it attempts to analyze the 'competition in innovation' dimension of the competitive process in much the same way as product market competition is assessed, by considering activity within an (innovation) market.

3.9. SOURCES OF INNOVATION

Innovation is associated with the learning process that results from the combination of both internal and external sources. It is useful to assess the extent to which sources of learning associated with innovation observed in the relevant market are external or internal to the firms in the market. As for practical purposes, we shall draw the distinction between:

(i) markets based on invention, where the sources of learning are essentially internal and dynamic change is largely due to firms' own creative and inventive actions, possibly in the form of R&D, or other sunk costs, investments or activities that lead to learning-by-doing,

(ii) markets based on adoption, where innovation derives largely from external sources such as suppliers or other sectors in the economy (research organizations).

We should accept that this distinction is not absolute. Each innovation, in every market, would generally involve both elements of invention and adoption. Few innovative improvements will be genuinely unique to the innovating firm since they will always involve a degree of imitation and adoption from observation of the world outside of the firm. Any invention will draw to some extent on past innovations and knowledge. Conversely, even where a firm seems to be directly adopting innovations associated with inventive effort by other organizations, it will need to adapt these to its own characteristics and circumstances, which requires an element of invention. Even where innovation is largely driven by inventive improvements in the services provided by suppliers, firms using such services may need some creative effort to incorporate these into their production processes and products. This view is underlined by comments from Nelson (1996, 244) following an extensive cross-country empirical study on innovation: "…the bulk of the effort innovation needs to be done by the firms themselves. While they may draw on outside developments, significant internal effort and skill is needed to complement and implement these."

Nonetheless, we can conceive of distinguishing the extent to which the innovation associated with supply of a future market is adoptive rather than innovative, even if both elements must be present to some degree.

In markets based on invention, a further distinction can be drawn on the basis of the sources of internal learning which has implications for the nature of the knowledge bases:

(i) R&D based markets. R&D is in many markets an important source of learning at the firm level both because it is associated with the production of new knowledge and because it allows the firm to access externally available knowledge.
(ii) In some other markets innovation is mainly driven by learning-by-doing. Learning results in a form of knowledge that is likely to be tacit, specific and relatively immobile, which implies that a firm would have to go down the learning curve without possibly relying on major knowledge spillovers from other firms.

An example of an adoptive market is a supplier-driven market where innovations are largely exogenous to the market and mainly embodied in equipment and components bought from other sectors. In these markets the process of innovation is essentially the process of diffusion of state-of-the-art capital goods and innovative intermediate inputs. The capability to supply in these markets requires essentially access to suppliers at terms that do not put it at a disadvantage relative to incumbent firms. The term 'suppliers' needs to be interpreted broadly to include both supplier firms and also organizations such as universities.

All else equal, innovation associated with adoption will yield a larger set of firms with capability to supply the future market which derives from that innovation.

Knowledge Base

Another important determinant of the ability to supply through innovation is the nature of the knowledge base on which innovation in the market is mainly built.

Dosi (1988) defines the notion of 'knowledge base' as:

> "…the set of information inputs, knowledge, and capabilities that inventors draw on when looking for innovative solutions."

For instance, Dosi suggests that in the case of microelectronics there are three major and complementary forms of knowledge: advances in solid-state physics; knowledge related to the construction of semiconductor manufacturing and testing equipment; and programming logics.

Technologies may differ in terms of the degrees of 'publicness' and 'universality' versus tacitness and specificity of their knowledge bases (Winter 1987). The tacit/codified nature of the knowledge base can be considered very important to determine the firms' capability to supply in a future product market:

All else equal, innovation associated with a codified, rather than tacit, knowledge base will yield a larger set of firms with capability to supply the future market which derives from that innovation.

Similarly for the general/specific nature of the knowledge base:

All else equal, innovation associated with a general, rather than specific, knowledge base will yield a far larger set of firms with capability to supply the future market which derives from that innovation.

This latter consideration suggests that an important element to consider is whether the knowledge base underlying innovation in the relevant markets is shared with other sectors in the economy. It may be the case that different sectors are similar in terms of the underlying knowledge base, so that a firm that operates in one would find it relatively easy to supply in the other market, with a similar knowledge base.

3.10. Adjunct Markets

An adjunct market can be defined, for current purposes, as one where the existence of demand in the market is dependent entirely on consumption in a related, but separate, relevant market (the primary market).

A commonly cited example is that of printers and compatible ink-refill cartridges, where demand for the latter is derived entirely from consumption of the former. In this case, the scope for competition in innovation directed to refill cartridges will be affected by whether cartridge manufacturers can interact successfully with the printers market.

A less obvious example is markets concerned with advertising (the supply of advertising space). In most cases, the existence of the advertising market is dependent on consumption in a related content market (be that the supply of television services, the supply of newspapers, magazines, or the supply of content on the Internet). Analysis of competition in the innovative activity related to entry in the advertising markets must pay attention to the need for the firm to simultaneously operate successfully in the content markets. So the relationship between advertising relevant markets and content relevant markets falls under the category of adjunct markets.

Of particular importance is where a competitive primary market is associated with an adjunct market in which there is monopoly. Thus firms compete in a broad primary market and each achieves monopoly supply in a narrow adjunct market (printers and refill cartridges). In such cases we may expect firms to use any marginal increase in profits realized in the secondary market to fund a decrease in price in the competitive primary market and thereby improve their relative position against competitors. Moreover, we would expect there to be checks against certain types of anti-competitive behavior in the monopoly linked markets due to the transfer of competitive effects from the primary market.

Thus in adjunct market cases, the effect may be such that innovative activity related to the secondary markets will depend on the ability of the firm to enter the primary market successfully; or the ability of the firm to ‘gain access’ to the necessary links between the primary and linked markets.

Network Effects

There seem to be clear similarities between adjunct markets and network effects, but the former is defined above as where consumption of a product in one market creates demand in another, while network effects are more related to how consumption of a product in one market increases demand for either (i) that product or (ii) a product in a different market that in turn may lead to increases in demand for the original product (see Gottinger (2003) for discussion of network effects).

Again network effects have important implications for the capability of firms to supply future markets associated with innovation. Supply to the future market will require a strategy for interacting with the network effects in such a way as to make the entry successful and sustain a position in the market. In particular, network effects can generate important benefits for incumbent technologies and thus create inter-temporal links between current and future markets, as discussed in the following section.

Intertemporal Dimension

Assessment of the capability to supply a future market may hinge dramatically on analysis of the 'intertemporal links' that exist between suppliers in a current market and potential suppliers to a future market. At a basic level there exist links over time between relevant markets wherever a firm's supply on a current market affects its capability to supply a future market, either positively or negatively. In many cases this is trivial, but cases may exist where the cumulative effects are so strong as to be the driving force of which firm will succeed in a distinct future market.

In particular, we would be interested in whether conditions are such that it is necessary to be competing strongly in a current market in order to have the capability to supply the future market, or whether the innovation associated with the supply of the future market is so drastic that entrants and incumbents have a similar capability to supply the future market. These factors will be key drivers of whether there is persistence of market power.

The links between different markets over time may be a critical factor in determining the ability of the process of competition to select the 'right' innovations from a new variety of products in a future market. If there are strong factors favoring persistence of leaders in current markets into future markets, this may undermine the benefits of competition in the innovation itself.

In many cases incumbency advantage should not be seen as detrimental. Consumers are likely to value similarities between current and future products, while producers will be able to make efficiency savings from use of similar production processes and reputation effects between current and future markets. But, to the extent that competition is possible in innovation and the supply of the future generation of products and technologies, this may bring sufficient benefits to warrant protection. The discussion of high profile competition cases shows (Hahn, 2003), even if there are benefits in only one operating system dominating current markets (e.g., due to economies of scale and customer familiarity) this does not preclude the benefits of competition to be the dominant supplier of 'next generation' operating systems.

Finally and very importantly, it should be possible to distinguish cases where any positive relationship between current and future markets is due to less-strategic characteristics of the market, such as its cost structure and nature of demand and the type of innovation associated with it; or the opportunity it provides incumbent firms to behave anti-competitively, in particular as regards competition in innovation.

Supply and Demand Side Links

The standard link on the supply side from one market to another comes simply from the fact that the production processes used to supply a current relevant market may be capable of use to supply a future market, even if the nature of the product supplied changes over time. This arises in particular where the same sunk costs are incurred to supply both current and future markets and therefore incumbents in the current market have an advantage to the extent that they have already incurred the costs.

In addition to standard production processes links, where the knowledge associated with the innovation into the future market is largely tacit, or non-imitable in another way, incumbents are in possession of a further advantage derived from the nature of the innovation itself. Conversely, where the innovation is more drastic and does not particularly build on tacit knowledge embodied in the supply of previous products we would not expect a strong advantage for current incumbents. Therefore analysis of the nature and source of innovation will feed into this analysis of inter-temporal effects.

When considering inter-temporal relationships, it is beneficial to understand that network effects can work through mechanisms based on both actual events and expectations. Most commonly, network effects are due to consumers at one period in time taking into account past consumption either by themselves or other consumers, when making their current purchase decision.

But network effects can also be generated by expectations. Consumers may form today expectations of what technologies and services other consumers will consume, and in turn lean towards products associated with this in order to benefit from network effects. Where this occurs firms are likely to be active in advertising and marketing their products to create an impression of high future consumption on current consumers, and to signal staying power in the market.

A further but perhaps less obvious type of (direct) network effect is economies of learning. Where consumers must invest in obtaining some specific knowledge and skills that facilitates consumption of the product, network effects may set in between products.

In the case of direct network effects, the relevant link between markets is actually directly between the current relevant market and the future market. In essence, network effects mean that on the demand-side the future market is linked to the current relevant market by the fact that consumption of a product in the current relevant market enhances demand for that product (or its descendant) in the future market.

In the case of indirect network effects, the relevant link between markets is indirect between the current relevant market and the future market, because it exists through a horizontally related market.

Such relationships are therefore important in a later stage, and discussed further below. The following table suggests possibilities for inter-temporal links between markets that may put firms which currently supply a related current market at a significant competitive advantage in the capability to supply the future market than firms that do not.

3.11. ACCESS SERVICES

Besides access to the innovation underlying the future market, a firm's ability to compete in the innovative activity related to a future market may also be dependent on key inputs that it must obtain from other firms and organizations. Such 'access services' describe the inputs that are needed to supply the future market, and therefore to compete in the innovation associated with that market.

Identification of access services and specification of the markets to which they pertain could be a valuable component of part of an analytical approach that can render analysis of capability to supply more explicit.

Several examples are explicit. A firm that wishes to supply television content to subscribers to a digital pay TV platform will need to obtain access services of some nature from the platform operator in order to distribute its content to viewers (unless it decides to sell the content wholesale to the platform operator). Or consider the supplier of a computer operating system. For independent software producers to supply products that are compatible with this operating system, these firms may require timely access to technical information that allows compatibility of software with the operating system. Regardless of how capable a firm is in the innovative activity relating to future computer software, its ability to compete in this regard depends on whether it can generate sufficient compatibility with operating systems, which may require licensing of some intellectual property rights from the operating systems supplier(s).

In dynamic markets, the existence of access services may be particularly important in determining market dynamics. Where a firm is dominant in the market for supply of an access service, it may be able to affect the evolution and development of the market(s) that the access services concern by preventing or otherwise impeding competition in the innovative activity that leads to the introduction of new products. In other words, access services represent a fundamental route through which a firm controlling the service would be able, unilaterally, to hinder the ability of rivals to supply a future market. In fact, in many cases analysis of potential exclusionary behavior can be made more robust and transparent by identifying formally the access services that pertains to this potential exclusion and then defining the market for this service and analyzing the competitive effects explicitly. This may be appropriate even if the access service is not seen as a traditional product in which there is trade.

An access service may be defined as a service that a firm must be able to procure (or generate itself in a more trivial case) in order to be able to effectively supply a related market. We do not relate the term to 'essential facilities doctrine' which in the context of EU Competition Law may have a more specific meaning based on certain criteria, although essential facilities would clearly be a sub-set of access services. Any necessary input can be seen as an access service, insofar as access to the input is necessary to supply downstream markets. But in some dynamic markets, the access service is perhaps less tangible than traditional intermediate input goods (e.g., it is a set of information protected by IP rights), and it will be particularly important to consider the existence of these services.

In some instances, there will be no current supply of the access service, particularly where the potential supplier of this service supplies the related market and is keen to maintain its current position within this related market. For instance, in the operating system and software example above, there is analytical value in defining the access services relating to the information (programming codes) that allow compatibility between platform and software, even if such information is not currently provided in the market.

3.12. Competition Policy Analysis

The previous tools on market definition and scope of innovation activities provide building blocks for the analysis of dynamic considerations in both antitrust and merger investigations. In an abuse of dominance for a dynamic market, the approach put forward confirms that market definition should remain a key element of the assessment of dominance and that capability to supply analysis can provide a structured approach to the analysis of possible exclusionary behavior and other anticompetitive practices relating to competition in innovation, and, in turn, competition on future product markets. The application of these tools to abuse of dominance cases is discussed next.

In a merger case, the analyses of market definition and capabilities to supply should provide the necessary building-blocks to assess whether the merger raises anti-competitive concerns both pertaining to competition in product markets and to competition in innovation. The application of these tools to the analysis of merger cases in dynamic markets, and similarly to agreements cases under Article 81 of EU Competition Law.

Before considering the application of the tools to these competition policy cases, we discuss the identification of some broad classes of economic environments on the basis of the distinction between the two dimensions of competition emphasized in the introduction to the study: competition on a product market and competition in innovative activity.

Competitive Constraints

In this type of environment there is no competition in innovation, but an expectation of product market competition once the outcome of the innovation is introduced (i.e., the outcome of the innovation will face competition from existing products). Although innovation is certainly a fundamental source of change of competition in the product market, nevertheless the absence of competition in innovation does not necessarily imply that the single innovating firm will be dominant once the innovation is introduced.

As discussed previously, market definition for future markets associated with innovative activity needs to be made with reference to whether the new product expected from successful innovation will face competition from the continued supply of current products in the future. This can exist for both horizontal and vertical technological change, and is most relevant where innovation is not expected to be drastic.

The temporal aspect of market definition may be important in determining this factor because in early stages of the life-cycle, the price of the new product may be so high as to entail a separate ‘niche’ relevant market for the new technology, which converges to a common relevant market with the old technology as relative prices become more reflective of relative quality differences implying scope for substitutability. Conversely, the new technology may start in the same market as the old, but the old may become increasingly obsolete and eventually drop out of the market.

Competition for the Market

Winner-takes-all markets are associated with cases where there is (often intense) competition in innovative activity but the future market is such that competition in it is, over a reasonable timeframe, not sustainable. Thus firms compete to attain a position of dominance.

Perhaps the most famous example of a winner-takes-all market is that for operating systems for desktop PCs or likewise laptops. It is instructive to recognize that this market benefits from massive economies of scale in production protected by IP rights (very low marginal cost of supply compared with very large fixed costs of initial product development) and substantial economies of scale in consumption (due in large part to the network effects associated with the relationships between the operating system market and the related applications software markets).

Identification of such markets is important because it affects the focus of competitive concerns.

In a winner-takes-all market, competition analysis should focus on whether a merger agreement by a dominant firm can harm competition in the innovative activity related to a future product market, rather than whether competition in the market will be restricted.

Most obviously, if there are strong grounds to believe that a future market is a winner-takes-all market, it is perhaps not appropriate for a competition supervision to block a merger or agreement between firms on the basis that this will create a dominant position or lessen competition in this future market. By definition, the nature of the market is such that its existence guarantees that a firm will be dominant on it, at least in the medium term. This illustrates an important point relevant to wider issues in competition policy: it is typically better to have a situation where a firm is dominant in a relevant market than for that market not to exist at all. Instead, any intervention must be based on the premise that the merger (agreement) lessens or distorts competition on some other, perhaps related market, or in competition in the innovative activity associated with the winner-takes-all market.

Similarly, in dominance cases, if we anticipate that a market is subject to winner-takes-all properties, then it is difficult to establish a case that a firm has abused its dominant position in monopolizing this market - the market is naturally prone to monopolization. Rather, analysis of an alleged abuse of dominance associated with this market should focus on how a dominant position in a related market (perhaps an access market) could be used to distort competition in the innovative activity associated with the winner-takes-all market, or how a dominant position in the winner-takes-all market could be abused to maintain that position — in effect used to distort competition in the innovative activity associated with the future generation of that market.

Of course, it is necessary to have a good idea of whether the future market in question is in fact a winner-takes-all market in which competition between products from the same generation of innovation is not sustainable. The box on the subsequent page discusses the difficulty of this assessment ex ante.

Winner-Takes-All Markets (WTAMs)

The key factors that would suggest WTAMs should have been identified in the analysis of capability to supply a future market because this should have identified reasons why 'winning' the race to the market would be associated with dominance, especially in terms of inter-temporal links.

WTAMs will arise for both demand and supply-side effects, and in practice both are likely to be required. Large economies of scale and network effects are likely to be the key drivers of this phenomenon: the question to ask is essentially whether the market can sustain competition between different products, or whether demand and supply side factors mean that competition can only take place through the next generation of innovations.

However, there does not seem to be a 'difference of kind' between such markets and those in which product market competition is sustainable, since in many cases consumers' heterogeneity in preferences, supported by attempts at product differentiation (physically or through marketing and branding) by suppliers, will allow a market to exist with a limited number of competing players even in the presence of network effects and large fixed costs.

Therefore, in most cases, it would seem inappropriate to conclude firmly that a future market will be a winner-takes-all market unless there is a comparable (and related) current market of similar characteristics displaying these features. Under this approach, analysis of whether future markets are likely to be winner- takes-all markets would focus on cases where current markets exhibit little product market competition, and seek to identify (i) whether there is scope for competition in innovation related to the supply of future generations of this market and (ii) whether the features of the current market that explain the lack of product market competition are likely to persist into the future generations of the market, even if competition in innovation took place.

Dynamic Persistence of Monopoly

Some markets may display no competitive pressures such that only one firm is engaged in the innovative activity associated with a future market, on which it is expected to face no competition.

Thus the products that are expected to result from the innovation are in separate relevant markets from current products supplied by other firms and there are no potential competitors to introduce the new products (i.e., the market is not winner-takes-all because there is only one firm in the race).

These markets are likely to exist in particular where there are such inter-temporal benefits in favor of an incumbent monopoly and where the innovation is not sufficiently drastic to allow new firms to undermine this advantage.

Note that where no current market exists to which the future market relates, i.e., the future market is entirely new, there is no persistence of dominance but simply an overall lack of competition in both innovation and (expected) supply dimensions.

3.13. Abuse of Dominance

We focus on the application of the tools to the abuse of dominance cases. In the Article 82 case of EU Competition Law the concern is whether the behavior of a firm (or a group of firms) in a dominant position (including collective dominance) impedes the competitive process.

While a firm operating in a dynamic environment may commit various types of abuse, the area most relevant to the analytical tools proposed in this study concern a dominant firm impeding other firms' ability to innovate and distorting competition in the supply of new or improved products in the future. This impediment will typically take the form of restricted access to a critical resource.

> "A dominant firm's actions can threaten dynamic competition if it monopolizes and restricts access to a critical resource that other firms need in order to innovate. The source of the dynamic problem is not the fact that the firm can earn monopoly rents but that it can restrict access, and so the appropriate remedy should focus on ensuring access." (Ellig, 2005, Chapter 9, 266)

This is not to say that the tools and guidance discussed in previous sections would be irrelevant to other types of abusive behavior. But the likelihood is that greatest value would come in cases where the allegation centers on whether a dominant firm has acted in a manner likely to distort competition in innovation and future markets.

Assessment of Dominance

For a firm (or set of firms) to be liable to current infringement of Article 82 it must be dominant within a current relevant market. Thus there is no need to define future markets to establish the applicability of Article 82.

The analytical guidance on market definition presented in Section 3.8 may be helpful in assessing the stability of the market. If we found that the current market definition was unlikely to be valid at some future point in time, this would provide useful information for the assessment of appropriate remedies. For example, if dominance is unlikely to persist, the value of certain undertakings on future conduct may be limited. But the fact that any dominance found is expected to be short-lived does not undermine the ability of a firm to abuse its dominant position while it lasts.

In addition, the guidance on market definition may be useful because the effect of the abuse could be felt in future markets. In particular, where the abuse of dominance is such as to restrict other firms' ability to innovate, the relevant area of impact is likely to be a future market.

Once the relevant markets have been defined, the remaining analysis consists in the analysis of dominance and the assessment of the effect of the alleged abuse on other firms' ability to innovate and supply a future market (and, where appropriate, development of remedies).

Since the concept of dominance applies equally to comparatively static and dynamic markets, we do not discuss the concept of dominance in any detail here. But among high profile antitrust cases, the Microsoft case as well discussed in the antitrust literature provides an economic interpretation of dominance that we would support for competition policy cases. Under such an interpretation, dominance would not be defined with any regard to whether firms can price above some hypothesized competitive price level (marginal cost or otherwise) or below in which case it would be related to 'predatory pricing' (McAfee, 2002, 208). Indeed a concept of dominance that relates to whether a firm can price above the competitive level treats competition as an outcome rather than a process; this seems particularly inappropriate for dynamic industries.

Instead, dominance in a relevant market is associated with cases where the constraints on a firm's current pricing (or quality, or innovation) of a product do not come from other products that can be reasonably considered competitors, but rather from goods and services in the economy as a whole. This description fits well with the legal definition of dominance that centers on the lack of competitive constraints on the behavior of a firm, which allows it to act independently of competitors and customers Note finally that, if a firm is dominant it is dominant in the supply of a service or set of services. Dominance does not apply at the firm level. Nor does it apply to innovation in abstract terms. If a firm seems to be 'dominant' in innovation this is because it is expected to be dominant in the supply of the output from that innovation – i.e., dominant on a future product market or dominant in the supply of an access service necessary for innovation by other firms. This supports the value of market definition, to define both current and future markets, and markets for access services.

After establishment of dominance the next step would be to uncover whether an abuse has been committed.

Assessing Market Abuse

In what concerns market dynamics, the application of the tools to assess the alleged abuse of dominance is centered on the observation that a dominant firm can distort competition if it can materially restrict other firms' capability to supply a future market. The capability to supply analysis seeks to provide a structured approach to the analysis of the capabilities and resources that a firm needs to successfully develop and introduce a product on the future market.

The analysis then involves an assessment of whether the conduct of the dominant firm has the effect of restricting other firms' capability to supply, rather than on whether this firm can completely preclude future supply (conduct does not have to completely foreclose competition and innovation to be abusive).

Perhaps most clearly, a firm may be able to distort competitors' innovative efforts by refusing to supply a key input needed by potential innovators, such as information relating to interoperability, a patent or a physical asset. This input would be identified in the final stage of the capability to supply analysis as an 'access service.' For example, the discussion of capability to supply previously commented on the case where competition in software innovation may be impeded if a platform operator refuses to provide independent software developers with the necessary information to develop new software products that will interoperate with the next generation of the platform product. In this case, the platform operator may be found dominant in the market for the (timely) supply of this information: the question for assessment for abuse would concern the extent to which the firm's refusal to supply this information could materially affect independent developers' capability to supply downstream software market.

The class of potential abuse extends wider, once we recognize that complete refusal to supply altogether is the extreme case of refusal to supply on reasonable terms. Even if the firm offers to supply the access service, the terms on which this offer is made could lead to distortions in innovation and competition. Examples of terms that may lead to impediments to innovation include: (i) discrimination between different (classes of) customers; (ii) bundling, where two products are only supplied as a single transaction; (iii) tying, where the price (or supply) of one service is conditional on purchase of subsequent units of that service, or of a related service; and (iv) other conditions of sales, such as exclusive purchase or supply.

We suggest that each of these types of potential abuse could be analyzed in the same manner (Evans et al. 2006, Chap.11). The fundamental issue to address is how the terms of trade under investigation affect other firms' capability to supply a future market, compared against a counterfactual of 'reasonable terms.'

Only by comparing the terms against a realistic counterfactual can we robustly assess whether it is the conduct in question that hinders innovation – and is therefore abusive — or whether some unrelated factor is affecting other firms' innovation.

Some predatory behavior would not fall into the categories of refusal to supply an 'access service' on reasonable terms, although it is sometimes very difficult to draw the distinction between conduct that is predatory and conduct that is exclusionary. Nonetheless, conduct, pricing or otherwise, that is intended to drive a competitor out of the market could also be detrimental to competition in innovation and on future products markets, as well as on current product markets. In this case, the capability to supply analysis can be used to assess the claim that particular conduct by a dominant firm, compared against a realistic counterfactual, can materially impede other firms' innovation and hence competition on a future market (Etro, 2008).

Finally, in considering the effect of the alleged abuse, the analysis can be structured in view of the resulting competitive effects. This should allow for a clear separation of effects of the conduct on the dimensions of competition in a relevant market and competition in innovation. This in turn may provide more robustness to the assessment of the particular innovation, in considering the effect of the alleged abuse.

As discussed above, if we anticipate that a market has winner-takes-all properties, then it is difficult to establish a case that a firm has abused its dominant position in monopolizing this market –the market is naturally prone to monopolization. Rather, analysis of an alleged abuse of dominance associated with this market should focus on how a dominant position in a related market (perhaps an access market) could be used to distort competition in the innovative activity associated with the winner-takes-all market, or how a dominant position in the winner-takes-all market could be abused to maintain that position – in effect used to distort competition in the innovative activity associated with the future generation of that market.

3.14. DOMINANCE AND MARKET DYNAMICS

The concept of dominance applies equally to what seem comparatively static and dynamic markets. The difficulties of dynamic cases may render it more necessary to assess dominance against an interpretation that is robust in economic terms. For completeness, we propose such an interpretation below. In respect of a case under Article 82, The European Court of Justice (ECJ) has held that

> "The dominant position referred to in Article [82] relates to a position of economic strength enjoyed by an undertaking which enables it to prevent effective competition being maintained on the relevant market by giving it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers."

It is instructive to consider in more detail how dominance relates to economic concepts of market power, and what relevance these have for market dynamics.

Simple economic models of perfect competition bring the result that perfectly competitive markets are characterized by firms setting price equal to marginal cost. This has led some to define market power as the ability of a firm to price above marginal cost. However, this result rests on a number of assumptions that are generally invalidated by the circumstances of the real world (in particular homogenous goods and full information). This does not undermine the value of any economic model that uses these assumptions. But it does mean that we should be careful when using such insight to analyze the nature and extent of competition in a market for the purposes of competition policy (or indeed any other form of economic regulation).

Dominance, as understood in EU case law, is a different concept to that of the ability of a firm to price above marginal cost. More fundamentally, competition policy does not even ascribe dominance to firms that can price significantly above marginal cost, as a reference to competition cases in the pharmaceuticals and media industries confirms. Thus competition policy practice clearly does not class firms as dominant or not dominant according to any marginal cost test (though, where it is appropriate, evidence on prices in relation to marginal cost may be used to support a finding of dominance). But if market power (dominance) in economic terms is not the ability to price above marginal cost, how is it to be defined? It is often argued that market power is the ability to price above the 'competitive level.' For example, antitrust agencies have commonly defined dominance with reference to the competitive price level.

Then, a dominant firm holding such market power would have the ability to set prices above the competitive level, to sell products of an inferior quality or to reduce its rate of innovation below the level that would exist in a competitive market.

We do not dispute here the value of this concept in abstract terms, but do call into question its value for practical purposes, especially in dynamic cases involving innovation.

It is possible to establish models of a ‘competitive price’ even in cases where such a price must be significantly above marginal cost to be consistent with competition existing (e.g., where firms must recover significant sunk costs).

Suppose it is possible to establish models of a ‘competitive price’ even in cases where such a price must be significantly above marginal cost to be consistent with competition existing (e.g., where firms must recover significant sunk costs). This is clearly appropriate for services such as the supply of patented pharmaceutical products or the supply of recorded music on CDs. It is no less true for services such as the supply of meals at restaurants.

But such analysis must accept that there is no clear way of establishing what rule should govern how each item of fixed costs should be allocated per meal out of the myriad of possibilities; that the competitive price may be different for peak-times and off-peak periods; that the firm’s current costs may not be indicative of the efficient levels; that current prices may reflect the need to pay a monopoly rental for premises; and that there are likely to be major inaccuracies inherent in such estimation. So long as there exist multiple plausible ways of calculating a competitive price, under different counterfactuals, methods and assumptions, there is simply no basis for using any one of these to define whether a firm is dominant. Thus, a practical definition of dominance cannot be based on the relationship between the price observed and some hypothetical competitive price level that must assume some arbitrary mark-up over marginal cost.

Moreover, it is not obvious that a firm that is dominant in a market (understood as one that is able to act substantially independently of competitors and consumers) would necessarily price above the ‘competitive level’ in that market. Suppose a firm serves two interrelated markets and is dominant in both. It might well be that the firm sets price below some notional ‘competitive level’ in one market in order to attract demand to the other market, and achieve maximum profits across the two. Therefore, in this particular instance, market power is associated with prices in one market being below the level that may be expected under a more competitive environment.

This point is even more important when we consider innovation. The discussion in the literature review has argued that in some cases an increase in competition may lead to more innovation (e.g., because firms may face greater risk of being excluded from a market if they do not maintain a high pace of innovation) while in others it may lead to less innovation (e.g., because a greater threat of imitation by rivals, once an innovation has been introduced, undermines the incentives for innovation in the first place). Therefore it does not make sense to define, or even describe, market power with reference to such power resulting in firms innovating at a reduced rate compared to a competitive counterfactual.

Thus we propose to abandon completely the view that dominance and market power, within the context of competition policy, are usefully related to any kind of price-cost margin or competitive benchmark.

Both practical economic reasoning, and reference to the ECJ definition of dominance suggest that dominance is associated with whether the constraints on the price of a service, or group of services, come from the supply (or potential supply) of other services that can reasonably be considered competitors, i.e., those of a similar characteristics and intended use.

Note finally that, in line with the arguments on market definition, if a firm is dominant it is dominant in the supply of a service or set or services. Dominance does not apply at the firm level. Nor does it apply to innovation in abstract terms. If a firm seems to be ‘dominant’ in innovation it is so because it is expected to be dominant in the supply of the output, or input, to that innovation.

3.15. Mergers and Acquisitions (M&As)

We discuss the application of the tools and guidance to competition policy analysis in the context of cases concerning mergers and cases concerning agreements between firms (under Article 81 EU Competition Law).

Both merger and acquisitions are associated with a reduction in the number of independently operating firms, and thus with the possibility of a loss of important competitive constraints between the firms involved. As the discussion below exposes, the underlying question in applying the tools to a merger investigation is whether the loss of competitive constraints (if any) is associated with the merger –which could be a loss of competitive constraints in current or future product markets and/or a loss of competitive constraints in innovation –significantly impedes competition in a manner that is not outweighed by pro-competitive effects.

Similarly, in investigation of an agreement between firms, the underlying economic implications are of the same nature: does the loss of competitive constraints (if any) associated with the agreement(s) between the parties significantly impede competition in a manner that is not outweighed by any pro-competitive effects? The key difference is in the examination of the exact nature of the agreement between the parties to establish the areas in which competition between the parties could be lost and the areas not affected by the agreement; this is not especially a dynamic issue and is therefore not addressed here.

M&A Issues

Mergers may have pro-competitive effects. Of particular relevance here, mergers may enhance the merged entities ability and incentives to innovate and supply a future market. Such benefits include the bringing together of complementary resources of the merging parties to enable successful innovation, the removal of wasteful duplication in R&D efforts and reduction in the risk of innovative activity by securing a more certain market for the outcome of the innovation.

Nonetheless, a merger, by removing competitive constraints that may otherwise exist between the merging parties, can harm the competitive process in terms of both competition in relevant products markets and competition in innovation.

Competition assessment in a merger case can be conducted by analyzing whether the loss of such constraints represents a significant impediment to competition, taking account of any possible pro-competitive effects of the merger, such as increased ability of the merged entity to innovate and supply a future market.

The approach to merger analysis proposed in this study explicitly distinguishes between competition in the relevant product markets and competition in innovation as a means to assess the impact of the merger on different sets of competitive constraints. For instance, the merger may result in the loss of competitive constraints in the product markets but not in innovation related to these product markets. In other cases, the merger may raise concerns with both competition in the market and competition in innovation. In general, the weights attributed to the consideration of these two dimensions of competition would depend on the specific circumstances of the case, e.g., the relative importance of innovation in the competitive process.

Merger Impact on Product Market Competition

If a merger creates a dominant position and impedes competition on a current relevant market, is this dominant position likely to remain in the future? Conversely, if a merger appears innocuous on current markets, is there still an impediment to competition in a future market?

An example of this is the analysis of a merger between two pharmaceutical companies that do not compete in any current relevant market but are investing in R&D to discover two drugs that, if successfully introduced, would compete in the future. Despite the absence of competition between the firms in current relevant markets, the merger may raise competition concerns that relate to future product markets. In this case, the state of competition in current product markets is completely uninformative of the state of competition in the foreseeable future, and the need arises to define a future market to allow competitive assessment. In the context of merger investigations, the definition of separate future relevant markets of this kind (i.e., market definition from the perspective of products yet to be introduced, rather than from the perspective of the future evolution of the current relevant markets) is most likely to be useful in cases where innovation is step-wise and the nature of future products can be reasonably anticipated in advance (such as products in a pharmaceuticals R&D ‘pipeline’).

At the other extreme, a merger may lead to the creation of a dominant position only on a current relative market. Whether this represents a significant impediment to competition may depend on how the current market definition is relevant into the future. This instability problem is perhaps most acute for those markets that, at the time of the investigation, are subject to significant changes in the main technological paradigm, especially when these changes are capability-destroying and thus likely to result in profound transformations of the competitive environment.

By maintaining the traditional approach to market definition based on the hypothetical monopoly test we apply this conceptual exercise to define both current and future relevant markets, within which competitive constraints can be assessed. This should render the analysis robust to the risk that the outcome of a competitive assessment of a product market would depend on the pointing time at which this market is considered. It would seem appropriate to discount the assessment of effects on competition in future markets to take account of the inherent analytic uncertainty pertaining to the future.

Effects of M&A on Competition in Innovation

In dynamic markets, innovation is often central to the competitive process and in some cases it is the only dimension on which competition takes place (e.g., winner-takes-all markets). Although the analysis of competition in M&A cases is normally focused on the assessment of competition in relevant markets, the introduction of some considerations pertaining to competition in innovation may be desirable, and in some cases necessary.

The approach proposed to complement the analysis of the effect of the merger on product market competition is based on the explicit consideration of competitive constraints that derive from competition in innovation. We might address the following issues: in which way would competition be impacted as a result of the merger? If so, would this result in an impediment to competition in innovation?

The focus at this first stage is on whether one party to the merger is undertaking innovative activity, or likely to undertake innovative activity in the future, that can be considered to exert a competitive constraint on the innovative activity of the other.

A next step is to define the future markets on which the output of the firms' innovative activity is expected to be supplied. These definitions cannot be entirely accurate; the analysis will be prospective, and will rely on realistic but uncertain assumptions as to the quality of the characteristics of the products developed, the marketing strategies adopted by suppliers and the preferences of consumers. Nonetheless, the best estimate of the likely future market (or markets) to which the innovative activity relates should provide valuable structure for assessment of the effect of the merger on competition in innovation. (This approach of estimating likely competitive constraints on the output of the innovation seems more robust than simply looking at firms' innovative inputs to assess competition.)

Once the relevant future markets have been defined it should be possible to analyze whether the merging parties are competing in innovation relating to the same future market(s). This can be undertaken by reference to the expected outcomes of the innovation and will therefore draw from some of the evidence used in the market definition analysis. If the two firms are found to be competing in innovation related to a future product market, this implies that the merger will result in the loss of a competitive constraint in innovation. If this is the case, the analysis will proceed to the next stage of analyzing whether the loss of this competitive constraint in innovation would significantly impede competition in innovation.

In an extreme case, there may be an innovation concern even if the merging firms are not currently engaged in competitive innovative activity. This would be so if there are reasons to believe that the merging firms are likely to compete in innovative activity in the foreseeable future. Such speculation would have to be very well-grounded if it were to affect the outcome of a merger investigation. Moreover, in practical cases, even where such a concern is valid, it is likely there would likely be other, more robust, reasons to block the merger.

In the alternative, if no future markets can be identified for which the merging parties are either engaged, or likely to be engaged in innovative activity to introduce products there would be little likelihood of an impediment to competition in innovation as a result of the merger.

Now would this result in an impediment to competition in innovation?

The focus at this second stage is on other firms with capability to exert a sufficient competitive constraint on the innovative activity of the merged party to compensate for the constraint on innovation removed by the merger.

In some cases it will be clear that the merger will not impede competition in innovation. For example, where many equally matched firms can be identified that are also competing in the same innovation, it should be straightforward to establish that the presence of these is sufficient to exert sufficient competitive pressures on the merged firm to prevent the merger impeding competition. However, in other cases it may be more complicated to assess the affect on competition in innovation. This may be so if only a very limited number of firms seem at first sight capable of competing in innovation with the merged entity. Or in cases where it is difficult to establish whether other firms' innovation would provide a sufficient constraint on that of the merged entity.

The purpose of the capability to supply analysis in this case is to identify those firms which have the capability to undertake innovation leading to supply of products on the future product market, within a reasonably close timeframe to the merging parties. These firms may currently be undertaking relevant R&D (e.g., accumulation of knowledge), in which case the analysis would seek to check that such firms have the other capabilities required to threaten the merging parties (e.g., access services and interaction with related markets). In other cases, a competitive constraint may be provided by the threat that firms could acquire the necessary knowledge (either internally or through the market) and have the incentives and other capabilities required. Finally, it should be checked that the constraint that would be provided by such firms on the innovation undertaken by the merging parties would be sufficient to compensate for the constraint removed by the merger, and thus prevent the emergence of a significant impediment to competition.

The overall effect of the merger on competition in innovation will also need to take into account the extent to which the merger itself will enhance the merged firms' capability to supply future markets. For example, where supply of a future market requires a specific access service, there may be strong benefits in consolidation between an innovator to the future market and a potential suppler of the access service. Thus the capability to supply analysis could be used to structure assessment of the significance of such an effect (although, as discussed below, any enhanced incentives for subsequent anticompetitive behavior would also require analysis).

Anticompetitive Behavior

Further to the direct effects on competition discussed above, a merger may also increase the risk of subsequent exclusionary (or predatory) behavior. This risk may arise if the merger leads to the creation of single or collective dominance (i.e., under Article 82). Or it might arise if the merger provides an already dominant firm with greater capability and motive for anti-competitive conduct.

In principle this is not simply a merger issue because, were the merger allowed to proceed, competition authorities would still be in a position to take action if such abuse were to take place. However, in practice, the competition authority may have legitimate reason to block a merger or seek undertakings in response to this risk. Such action may be appropriate if the authority anticipates significant difficulty in observing or proving abuse of Article 82, or implementing sufficient remedies, once the abuse has occurred. Since such concerns may well relate to competition in innovation, and on future markets, the guidance provided for application of the tools to Article 82 cases would apply here, albeit in a prospective manner.

References

Aghion, P. & Griffith, R. (2005). Competition and Growth, Cambridge, Ma.: MIT Press.

Arrow, K. J. (1974). "Limited Knowledge and Economic Analysis," *American Economic Review*, *64*(1), 1-10.

Besanko, D., Dranove, D. & Shanley, M. (2000). 2nd edition, Economics of Strategy, New York: John Wiley&Sons.

Brenner, R. (1987). Rivalry in Business, Science, Among Nations, Cambridge: Cambridge Univ. Press.

Case, J. (2008). Competition: The Birth of a New Science, New York: Hill and Wang.

Davis, P. & Garces, E. (2010). Quantitative Techniques for Competition and Antitrust, Princeton, N. J.: Princeton Univ. Press.

Dosi, C. (1988). "Sources, Procedures and Macroeconomic Effects of Innovation," *Journal of Economic Literature*, *26*(3), 1120-1171.

Foray, D. (2006). The Economics of Knowledge, Cambridge, Ma: MIT Press.

Easley, D. & Kleinberg, J. (2010). Networks, Crowds and Markets, Cambridge: Cambridge Univ. Press.

Ellig, J. (2005). ed., Dynamic Competition and Public Policy: Technology, Innovation and Antitrust Issues, Cambridge: Cambridge Univ. Press.

Etro, F. (2008). Competition, Innovation and Antitrust: A Theory of Market Leaders and its Policy Implications, Berlin: Springer.

Evans, D., Hagiu, A. & Schmalensee, R. (2006). Invisible Engines, How Software Platforms Drive Innovation and Transform Industries, Cambridge, Ma.: MIT Press.

Fisher, F. M., McGowan, J. J. & Greenwood, J. E. (1983), Folded, Spindled, and Mutilated, Economic Analysis and U. S. v. IBM, Cambridge, Ma.: MIT Press.

Gottinger, H. W. (2003). Economies of Network Industries, London: Routledge.

Gottinger, H. W. (2006). Innovation, Technology and Hypercompetition, London: Routledge.

Gottinger, H. W., Umali, C. & Floether, F. (2010). eds., Strategic Alliances in Biotechnology and Pharmaceuticals, New York: Nova Science.

Grossman, G. M. & Helpman, E. (1993). Innovation and Growth in the Global Economy, Cambridge, Ma: MIT Press.

Greaseley, D. & Oxley, L. (2011). Economics and History: Surveys in Cliometrics, Oxford: Wiley-Blackwell.

Hahn, R. W., ed. (2003). High-Stakes Antitrust, The Last Hurrah?, Washington, D.C.: AEI-Brookings Joint Center for Regulatory Studies.

Jacquemin, A. (1987). The New Industrial Organization, Market Forces and Strategic Behavior, Cambridge, Ma.: MIT Press.

Kirzner, I. (1973). Competition and Entrepreneurship, Chicago: Univ. of Chicago Press.

Kwoka, J. E. & White, L. J. (1999). 3rd edition, The Antitrust Revolution, New York: Oxford Univ. Press.

Lerner, J. (2012). The Architecture of Innovation, The Economics of Creative Organizations, Boston: Harvard Business Review Press.

Marshall, R. C. & Marx, L. M. (2012). The Economics of Collusion, Cambridge, Ma.: MIT Press.

McAfee, R. P. (2002). Competitive Solutions, Princeton, N. J: Princeton Univ. Press.

Motta, M. (2004). Competition Policy, Theory and Practice, Cambridge: Cambridge Univ. Press.

Nelson, R. R. (1996). "The Sources of Growth," Chap.10 in National Innovation Systems, Cambridge, Ma.: Harvard Univ. Press, 274-301.

Nouel, G. L. (2001). "Competition Assessment of Vertical Mergers and Vertical Agreements in the New Economy," Brussels: EC Commission, Nov., http://europa.eu.int/comm./enterprise/lib_competition/doc/merger_agreement_Study pdf.

Porter, M. E. (1996). On Competition, Boston: A Harvard Business Review Book.

Shapiro, C. & Varian, H. (1999). Information Rules – A Strategic Guide to the Network Economy, Boston: Harvard Business School Press.

Shelanski, H. A. & Sidak, J. G. (2000). "Antitrust Divestiture in Network Industries," *The University of Chicago Law Review, 68*, 95-146.

Simon, H. (2009). Hidden Champions of the 21st Century, New York: Springer.

Spulber, D. F. & Yoo, C. S. (2009). Networks in Telecommunications, Cambridge: Cambridge Univ. Press.

Stigler, G. (1968). The Organization of Industry, Chicago: Univ. of Chicago Press.

Tirole, J. (1988). The Theory of Industrial Organization, Cambridge, Ma.: MIT Press.

Tirole, J. & Laffont, J. J. (2001). Competition in Telecommunications, Cambridge, Ma.: MIT Press.

USDOJ (1995). IP Guidelines, Dept. of Justice (DOJ), Washington, D.C.

Whinston, M. D. (2006). Lectures on Antitrust Economics, Cambridge, Ma: MIT Press.

Winter, S. (1987). "Knowledge and Competence in Strategic Assets," in Teece, D. ed. The Competitive Challenge, New York: Harper and Row.

Chapter 4

INCREASING RETURNS MECHANISM

Increasing returns mechanisms are likely to emerge through networks, competition, innovation, new technology and dynamic industrial complementarity as outlined by Roberts (2004, 34-51) for the modern firm. Increasing returns also characterize the internal economy of large firms and corporations and their supply-chain network (Radner,1986). For an industrial policy toward promoting economic growth it would be of prior importance to complementarize Increasing Returns Industries (IRIs) which would enhance their value creation and competitiveness. Competition in innovation is likely to lead to IRIs triggering serial innovations inducing more IRIs.

Most industrial sectors of advanced industrialized economies are not perfectly competitive. They are usually formed by a small number of big firms with non-negligible market share. Besides being prevalent in the economy, big firms cluster around concentrated industrial structures which exhibit a skewed distribution of firm size and market share. This situation may be brought about by the intrinsic potential of dynamic technological competition to end up in (temporary) technological monopoly, so in those cases industrial competition may start out symmetric but end up asymmetric.

A prevalence of IRIs make a sharp distinction among the growth of the firm, industrial structure and competition between technologies artificial. Competition between firms and competition between technologies are very closely related. Since as much as technological standards tend to be monopolistic, they favor the emergence and persistence of dominant firms and concentrated industrial structures. And conversely: factors causing the growth of the firm and industrial concentration may accelerate the adoption of a technology. Granted that business competition is a process where technological competition and the growth of the firm intermingle and cause each other, then a situation where firm dominance and rivalry in new technologies compete is complementary to each other, thus it should form a comprehensive whole.

We show how the competitive process proliferates in IRIs where the total of all unit activities linked together yield a higher return than the sum of the individual unit activities operating separately.

The purpose of this review is to show that we cannot accurately understand industrial competition without taking into account the self-reinforcing nature of commercial success in most emerging markets. We enrich increasing returns mechanisms by incorporating a set of stronger, yet neglected, increasing returns mechanisms --reputation effects, infrastructure effects and positive network externalities into an interactive framework model. The resulting theoretical structure, we will argue, captures the interdependent and cumulative character of the three aspects of industrial competition: the number and size of firms, skewed industrial structures, and the nature of technological competition.

The increasing returns discussion in economics has provided important insights into the characteristics and dynamics of modern industrial economies. However, the discussion on policy applications had some misleading features in the past to conclude that a completely new network economy is emerging and that it obeys a set of rules, which are totally different from those that apply to traditional sectors of the economy. While it is undeniable that the increasing returns paradigm delivers new insights and while there is no doubt that this paradigm is key to our understanding of new industrial sectors, and their sustaining role in productivity growth, we should clarify its proper role in industrial structure and growth of the economy. At this stage we are most concerned about the catalytic role of technological competition in increasing returns industries. Increasing returns industries are nowadays most likely to be identified with high technology industries, in particular with information, communication and health care related industries (Gottinger, 2003).

As an example, in a corporate context, how to unlock increasing returns in its global operations, consider General Electric (GE). It constantly evolves its portfolio to drive growth despite its large size and already significant presence in major markets. Under its (former) CEO Jack Welch it encourages its executives and business units to take an expansive view of its markets as a means of unlocking growth initiatives that a product centric view would miss. Often, when its market share exceeds ten percent, it seeks to redefine the market more broadly to include adjacent products or services. This continual questing lies behind successful moves from manufacturing to services, that has allowed it to keep growing in complementarizing given industrial markets (This even remains true, as in recent financial distress situations where complementarizing growth activities could accelerate a downward cycle).

4.1. Supply-Side and Demand-Side Scale Economies

A first source of increasing returns assuming constant technology identifies a concentrated industry structure as a result of supply-side scale economies. In many cases large firms are more efficient than smaller companies because of its scale: larger corporations tend to have lower unit costs. This efficiency in turn fuels further their growth. However, positive feedbacks based on supply-side economies of scale usually run into natural limits. Past a certain size companies find growth difficult owing to the increasing complexity of managing a large organizational structure. From then on, negative feedback takes over. As traditional supply-side economies of scale generally become exhausted at a scale well below total market dominance, large firms, burdened with high costs, never grow to take the entire market and smaller, more nimble firms can find profitable niches. Shapiro and Varian (1999) conclude that because of this most industrial markets are oligopolies rather than monopolies.

Negative feedback generated by the difficulties of managing large organizations (scale diseconomies) indeed interrupts the growth of the firm and the level of industrial concentration. This situation, nevertheless, may be transient, because firms may be subject to other sources of increasing returns. Large firms that go through increasing returns mechanisms other than scale economies may increase their efficiency and overcome the negative aspects of overgrown organizations. Industries in which scale diseconomies are counterbalanced by other increasing returns mechanisms, then, may begin to head toward the extreme of ‘winner-takes-most’ situation. The increasing returns mechanisms capable to offset scale diseconomies are usually related to technological progress, so in the following sections we analyze other major causes of the growth of the firm, namely, the Schumpeterian loop, cost reducing learning, learning-by-doing, learning-by-using, and demand-side increasing returns.

Demand-side scale economies is a new and virulent form of positive feedback. Unlike supply-side economies of scale which reach a limit when the market grows, they have no limit. Common information economy products with many compatible users are usually subject to this kind of positive feedback because they become more and more valuable to each user as they are bought by ever more users. The combination of the two even results in an extraordinary industrial impact in which growth on the demand side both reduces cost on the supply side and makes the product more attractive to other users − accelerating the growth in demand even more. The result is especially a strong positive feedback, causing entire industries to be created or eliminated far more rapidly than through organic growth or decline during the industrial age.

4.2. Schumpeterian Mechanism

The most widely accepted theory of technological change in modern economics is Schumpeter’s (1942). In the Schumpeterian world, scale economies are present as well, but technology is not a constant. Here the creative role of the entrepreneur, endogeneously driven by ‘animal spirits’ allows for the introduction of new technologies capable to displacing the established ones (Francois and Lloyd-Ellis, 2002). In the context of IRIs, Schumpeter’s main point − through subsequent interpretation −has been that innovation competition leads to increasing returns economies triggering serial innovations inducing more IRIs (Freeman, 2003). Most of Schumpeter's discussion stresses the advantages of concentrated market structures involving large firms with considerable market share. In his view, it is more probable that the necessary scale economies in R&D to develop new technologies be achieved by a monopolist or by the few large firms of a concentrated industry. Large size firms, besides, may increase their rate of innovation by reducing the speed at which their transient rents and entrepreneurial advantage are eroded away by imitators. In the absence of patent protection large firms may exploit their innovations on a large scale over relatively short periods of time −and in this way avoid rapid imitation by competitors − by deploying their productive, marketing and financial capabilities. Large firms may also expand their rate of innovation by imitating and commercializing other firms’ technologies.

Schumpeter’s thesis encouraged a large body of empirical literature in the field of industrial organization. Most of this literature focused on two hypotheses associated with Schumpeter's assertion: (1) innovation increases more than proportionally with firm size and (2) innovation increases with market concentration.

The most recent comprehensive review of the empirical evidence of the relationship between innovation and firm size and market structure is due to Cohen and Levin (1989). They observe that the empirical results on the Schumpeterian relation are accurately described as fragile. They note that the lack of robust results seems to arise in part from the inappropriate attention to the dependence of these relationships on more fundamental conditions. From their overview Cohen and Levin draw the basic methodological lesson that the omission of important and potentially correlated variables that influence innovation can lead to misleading inferences concerning firm size and concentration.

Following Schumpeter's lead, Nelson and Winter (1982) stand out for having formalized and completed many of Schumpeter's original intuitions. Whereas the connection between industrial structure and innovation has been viewed by Schumpeter as going primarily from the former to the latter, in Nelson and Winter (1982) there is a reverse causal flow, too. That is, there is clearly a circular causality suggesting a self-reinforcing mechanism between the innovations and the firm's growth. Nelson and Winter (1982) stand out not only for having recognized the endogeneous character of innovation and market structure, but also for having pointed out and modeled the mutual causality between technical change and market structure (Nelson, 1986).

Evolutionary economists (like Nelson and Winter) define innovation very broadly. It encompasses product and process innovation, opening up new markets, and acquisition of new sources of raw material. They also describe the nature of technical progress as succession of major discontinuities detached from the past and with a quite transitory life span. This process of change is characteristic of certain industries, but it is not the sole kind of technological change. Technological change can also be *continuous.* That is to say, technologies improve constantly in absolute terms after their introduction. The view of technological progress as a continuing, steady accumulation of inumerable minor improvements and modifications, a routinization of innovation, with only very infrequent major innovations, has two sources: (1) the accumulation of knowledge that makes possible to produce a greater volume of output from a given amount of resources and (2) the accumulation of knowledge that allows the production of a qualitatively superior output from a given amount of resources. The former source of technological progress is the result of a *cost reducing learning process,* while the second category is the result of what is known as *learning-by-doing* and *learning-by-using.* Given that both categories of technological progress are important determinants of the number and size of firms in a given industry, we analyze them in the next sections.

Cost Reducing Learning

An important aspect of technological change is cost reducing in nature. In strategic economics, Porter (1980) and Henderson (1975), pioneered the notion of experience curve as a source of cost reductions. Also in economics, Hirsch (1956) has underlined the importance of repetitive manufacturing operations as a way of reducing direct labor requirements, while Arrow (1962) has explored the consequences of learning-by-doing (measured by the cumulative gross investment, which produces a steady rate of growth in productivity) on profits, investment, and economic growth. However, the historical study on the patterns of growth and competitiveness of large corporations of Alfred D. Chandler (1990) is a major and detailed contribution to our understanding of the way firms grow by diminishing costs.

Large corporations, according to Chandler, along with the few challengers that subsequently enter the industry, do not compete primarily on the basis of price. Instead they compete for market share and profits through functional and strategic effectiveness. (Caminal and Xives, 1996). They compete functionally by improving their products, their processes of production, their marketing, their purchasing, and their labor relations. Big corporations compete strategically by moving into growing markets more rapidly and effectively than do their competitors. Such rivalry for market share and profits make more effective the enterprise's functional and strategic capabilities, which, in turn, provide the internal dynamics for continuing growth of the enterprise. In particular, it stimulates its owners and managers to expand into distant markets in its own country and then to become multinational by moving abroad. It also encourages the firm to diversify and become multiproduct by developing and introducing products in markets other than the original ones.

Learning-by-Doing

Some of the writings on industrial competition assumes that firms compete mainly in cost-reducing competitive advantages, especially those achieved through scale economies, scope economies (economies of joint production and distribution), and innovation in production and organizational processes. Here technical progress is implicitly treated as the introduction of new processes that reduce costs of producing essentially unchanging products. Beyond, there is a category of learning known as ‘learning-by-doing’ (Rosenberg, 1982) which enhances the qualitative aspects of final products.

Western industrial societies today, Rosenberg (1982) argues, enjoy a higher level of material welfare not merely because they consume larger per capita amounts of the goods available. They have also made available improving forms of rapid transportation, instant communication, powerful energy sources, life-saving and pain-reducing medications, and other goods that were undreamed of one or two centuries ago. Therefore, ignoring incremental product innovation and quality improvements in products is to overlook what has been one of the most important long-term contributions of technical progress to human welfare. Many products, such as beverages, toothpaste, soap, clothing, VCRs, TV sets can be subject to improvements. Such improvements, however, are marginal when compared with the amazing rate of development that other products and technologies can reach. Automobiles, aircraft, flight simulators, computers, and nuclear reactors are very complex technologies and, as a consequence of this, have a tremendous capacity of being enhanced. Consequently, the competitive behavior of the firms that produce these technologies consists not only of the innovative acts they perform to improve production, organizational, and distribution processes, but also from the efforts to improve constantly their products.

Learning-by-Using

With respect to a given product, Rosenberg (1982) distinguishes between that kind of learning that is internal to the production process (learning-by-doing) and that which is generated as a result of subsequent use of that product (learning-by-using). The later category of learning begins only after a certain new product is used. In an economy where complex new technologies are common, there are essential aspects of learning that are a function not of the experience involved in producing a product but of its *use* by the final consumer.

The optimal performance of durable goods (especially complex systems of interacting components) often is achieved only after intensive and prolonged use. In the aircraft industry, for instance, the attainment of high standards of reliability is a major concern, in particular during the development stage. But it is only through extensive use of aircraft by airlines that faults are discovered and eliminated and detailed knowledge is gained about metal fatigue, weight capacity, fuel consumption of engines, fuselage durability, minimum servicing, overhaul requirements, maintenance costs, and so on.

4.3. Demand Side Increasing Returns

As indicated in Sect. 4.1 in the economy there are increasing returns mechanisms that come from the demand side of the market, not just from supply side. For the average (risk adverse and imperfectly informed) consumer it becomes more attractive to adopt a widespread technology or product. Minimizing the risk of purchasing a defective technology or the cost of searching for an adequate one introduces a reputation or informational feedback that may produce a disproportionately high selection of the best-selling option. Informational or reputational feedback effects occur in various situations that could be reinforced through network externalities. First, when the complexity of the technology or product in question is such that consumers try to reduce uncertainty by asking previous purchasers about their experience with these technologies (Arthur and Lane, 1993). Second, in other situations the source of uncertainty is not the complexity of the technology, but the large quantity of options the consumers face. One is bound to choose, and the best way to do so is by confining one's attention to the best-assessed items in the consumer report. Third, in a market where the quality or value of a product is defined on the basis of arbitrary and short-living conventions, rather than strictly on the basis of lasting objective value, consumers usually tend to follow the expert's opinion. This kind of easy-to-manipulate, reputation-driven market success is typical of markets for highly symbolic products (e.g., many art markets, fashion wear and luxury items), which also will result in a disproportionately high selection of one of the options.

Finally, the most preeminent and common kind of reputation effects in the economy, arise plainly as a result of a well-timed and very aggressive advertising campaign. This self-reinforcing mechanism -and the lasting market dominance that it causes- might be quite unrelated to relative added value, but it certainly might produce an excessive predilection for one of the options.

By moving beyond the Schumpeterian hypotheses and focus on a more complete model of industrial competition we have identified other fundamental determinants of technological change that affect the mutual link between firm size and market structure (Aghion and Howitt, 1998).

These determinants – which in our analysis take the form of increasing returns mechanisms –are usually studied as if they work independently from the other. But there are not many cases of industries where one single mechanism acts in isolation from the other sources of increasing returns. Therefore, the growth of the firm and the evolution of skewed industrial structure, more than the result of a single self-reinforcing mechanism, are the effect of the combination of several sources of increasing returns, which overlap and feed back upon one another.

The unification of the resource-based loop, the Schumpeterian loop, scale economies, the different categories of learning, and demand-side increasing returns (reputation) – loops A, B, and C, respectively, in Figure 4.1– constitutes a simple but useful model capable to explain endogeneously the number and growth of firms in a given industry, and in a wider context, the gap of economic performance in a given industrial sector.

In the model sketched in Figure 4.1 the positive relationship that runs from industrial structure to efficiency operates through the accumulation of rare resources, innovations, scale economies, reputation, and the different aspects of learning. This dynamics, over time, makes costs fall as learning accumulates, new technologies are developed and improved, and firm-specific factors are amassed and exploited due to output increases. As a result of this mutual causality, market share and production levels increase, price falls, profitability rises, and with which relatively profitable firms expand continually while unprofitable ones contract uninterruptedly. Ricardian rents qualify as economic rents assigned to legal market opportunities rather than a category of rent-seeking (Baumol, 1993, Chapter 3).

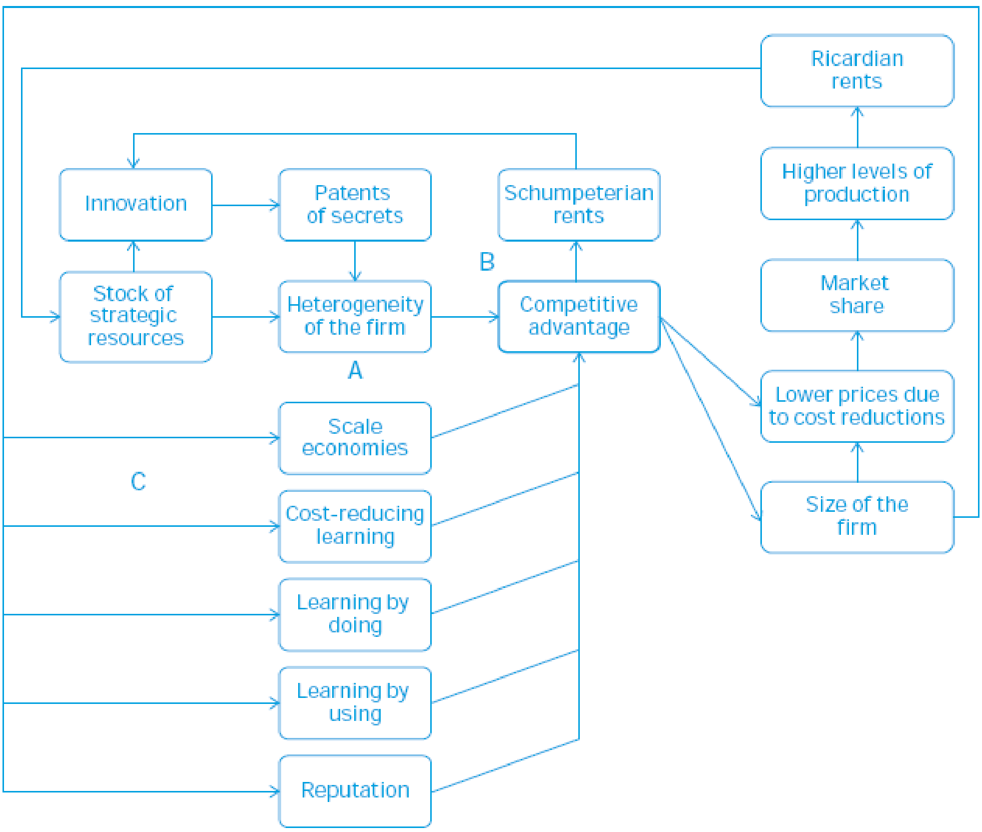

Figure 4.1. Increasing Returns Mechanism: A Qualitative Model of Industrial Competition.

A relevant aspect of the structural determinants of the number and size of firms in an industry suggested in this model is that, when one of them is exhausted, causing a slowdown in the growth of the firm, the other mechanisms may be activated, which may allow for a further period of continued rapid growth. When the firms of a given industry are capable to accumulate firm-specific resources, innovations, cost reducing learning, qualitative product innovation based on learning-by-doing and learning-by-using, and reputation, these firms usually use them as strategic weapon. In doing so, they are capable not only to neutralize but also to overwhelm the negative effects of complex, overgrown, hard-to-manage organizational structures that arise from their constant growth. The process can take a long period of time, but eventually the sources of increasing returns can drive markets toward increasingly skewed industrial structures. For instance, in the commercial aircraft industry competition principally involves considerable development costs, continuous improvements in aircraft models, technology and product support, so this industry exhibits substantial scale economies, large scope for learning-by-doing, learning-by-using, and reputation effects. Because of this, the commercial aircraft industry has been characterized by an increasing skewed industrial structure. The structure of this industry, after the acquisition of McDonnell Douglas by Boeing, was virtually reduced to a monopoly in the United States. In the world aircraft market Boeing only competes with EADS, Airbus Industries. It is obvious that the merger of the two main manufacturers of the American aircraft industry should have brought about some gain in efficiency, which counterbalanced the diseconomies owing to the managing a more complex organization. Otherwise, the merger would not have taken place or would have been the result of irrational behavior. The structure of some industries does not seem to head toward monopoly. However, over time, their level of concentration has increased substantially. The world automobile industry, for instance, in 1970 was composed of at least 40 major competitors. In 1998, with some mergers and acquisitions, the number of the main manufacturer was reduced to 17. Because of large possibilities to accumulate cost reducing learning and the large scope for qualitative product improvements in the world automobile industry, both the number and the homogeneity of the firms competing in this industry are expected to decrease even further in the future (Clark and Fujimoto, 1991) Here, again, benefits due to both cost reducing learning and qualitative product innovations brought about by mergers and acquisitions are larger than any cost created by scale diseconomies. Another interesting aspect of this model is that it also offers an endogenous explanation of the number and size of firms. In contrasts with the traditional economic views---that see industrial structure (number of firms) as an exogenous variable and assume homogenous firms--- and the strategic paradigms---which are focused first and foremost in explaining heterogeneity among firms within an industry-, this model recognizes that the strategic choices and acts of the firms have an effect not only on the performance and size of the firm itself, but also on the structure of market. In summary, industrial structure is caused by a combination of various increasing returns mechanisms. Here, then, the combination of accumulation of resources, product innovation, scale economies, cost reducing learning, learning-by-doing, learning-by-using or reputation enhances the performance of the firm and determine, to a great extent, the level of skewness of the structure of the industry where it competes.

4.4. INCREASING RETURNS AND ERGODIC MARKETS

Traditional neoclassical economics has tended to portray most economic situations as something analogous to a large Newtonian system, with a unique and stable equilibrium solution predetermined by a given pattern of resources, preferences, and technological possibilities. Brian Arthur and his group (Arthur, 1994; Arthur et al., 1987), however, have shown that this conventional way of seeing economic reality overlooks important and frequent economic situations where increasing returns are conspicuous. In order to distinguish economic situations characterized by decreasing returns from those where increasing returns are dominant, Arthur, Ermilov and Kaniovski (1987) developed IRM specific applications to the theory of *non-linear Polya processes*, which describe the long-run self-organizing structures that emerge from dynamic processes where proportions are involved. The general non-linear Polya scheme can be pictured by imaging an urn of infinite capacity to which balls of several colors are added. In the simplest case, where decreasing and constant returns prevail, the probability of a ball of a given color to be chosen the next time is independent of proportions of colors at the moment of the addition. In this simple sequential process, the strong law of large numbers predicts that, over time, the proportion of balls of color i has a fixed probability, where $\Sigma_i q(i) = 1$. Therefore, it has a unique, predetermined outcome. Sequences of choices in these simple cases are important at the beginning of the process. However, as the process advances, different sequences of choices are averaged away by the economic forces, which are subject to constant or decreasing returns. So, no matter the sequences of choices, the system will always – with probability one – end up into the same pattern. For instance, in a coin-tossing experiment the event ‘head’ is independent of previous tosses, then the expectation of a ‘head’ in each toss is 0.5 no matter how many times the experiment is repeated. Likewise, the proportion of 6's in a dice-casting experiment will tend to 1/6. The process by which firms in an industry concentrate in different regions is like the coin-tossing or the dice-casting experiment, if the geographic preferences of each firm is not modified by the preferences of the other firms. In more general cases -where increasing returns are present–the dynamics is completely different and the standard strong law is inapplicable. In this regime, the next ball to be added into the urn is not known, but the probability of adding one ball of specific color depends on the present proportions of colours in the urn. In other words, the probability of an addition of the colors becomes a function of the proportions of ball of each color at each time of choice. The case of firms deciding where to settle down illustrate this kind of non-linear Polya processes. Here increasing returns can be incorporated within the model by introducing agglomeration effects. Because of agglomeration effects, additions to a specific region are not independent of previous locational choices and firms are added incrementally to regions with probability exactly equal to the proportions of firms in each region at the time. Under increasing returns, then, the process becomes *path-dependent.*

Arthur et al. (1987) show that at the outset of the process proportions are not stable, but once the industry settles into a vector of proportions, locational patterns become constant at that vector with probability one. However, the constant vector is selected randomly from uniform distribution over all possible shares that sum to 1.0. This means that each time this locational process is rerun under different historical events, it will in all likelihood settle into a different pattern. Therefore, it is possible to predict that the locational pattern will tend toward a constant proportion, but it cannot be foreseen at *which* proportion it will settle down.

The interpretation of economic history is different under different regimes. Under constant and diminishing returns, the evolution of the system is *ergodic.* Ergodic structures emerge when repeated random events -that are drawn from the same distribution and are independent from previous ones- have a long-term average that approach their expected value, that is a long-run persistent random cycle (Feller, 1957, 410). The typical example of an ergodic system is coin tossing. If a fair coin is tossed indefinitely, the proportions of heads varies considerably at the outset, but settles down to 50 per cent with probability one. The evolution of an ergodic system, therefore, follows *a convex* probability function, which has expected motions that lead toward a unique, determinate outcome. In this regime 'historical chance' cannot influence the evolution of the systems so history is reduced to a mere deliverer of the inevitable and the known.

Under increasing returns, by contrasts, the process is *nonergodic,* because small historical events become magnified by positive feedback. A nonergodic system follows a nonconvex probability function, so two or more outcomes are possible and 'historical chance' determines which of these is ultimately selected. History becomes all-important. There are some cases of nonergodic systems in which, from the multiplicity of structures that may emerge, there are some 'corner solutions' with a single option monopolizing the choices. In this specific kind of nonergodic systems, while information on preferences, endowments and transformation possibilities allows locating and describing the various possible corner equilibria, it is usually insufficient to determine which one will be selected. In these cases, as Arthur (1994, 13) has pointed out 'there is an indeterminacy of outcome.'

Adoption of technologies that compete under diverse regimes can be appropriately modeled as a nonlinear Polya process (cf. Arthur, 1989, 1990a, 1990b; Arthur et al., 1987). In the simplest regime, when technological competition is characterized by constant and decreasing returns, the probability of a technology of being chosen depends on its current market share. As each adoption is independent of the previous one, market share should converge to a point where they equal probability. Therefore, under constant and decreasing returns two technologies or products performing the same function will end up sharing the market according to each technology's intrinsic value and technical possibilities. Therefore, markets characterized by constant and decreasing returns can be called ergodic markets.

Under increasing returns to technological adoption, the probability of adoption depends on the numbers of adoptions holding each technology at a particular time.. Markets of this kind can be called nonergodic markets. Within this kind of markets there are those where increasing returns may drive the outcome toward a single dominant technology, with small events early on selecting the technology that takes over. This particular type of nonergodic markets can be termed tipping or indeterminate markets. This indeterminacy relates to the 'selection problem' – how one allocation outcome is 'selected' over time by small historical events when there are several possible long-term results. Thus non-ergodicity in dynamic physical systems appears to be equivalent to historical 'path dependence' in economic systems (Arthur, 1999).

4.5. Technological Competition under Uncertainty and Inertia

In high technology markets the commercial success of emerging new technologies is both highly uncertain and inertial. As regards uncertainty, in addition to the problem trying to discern the true potential of a new generic technology, there is also the difficulty of foreseeing the precise direction in which the said technology will evolve. Indeed, as depicted in Figure 4.2, the emergence of a new generic technology generally opens the door not just to one specific technological path, but rather to a whole variety of possible trajectories in product design and process technology.

On the other hand, the inertial forces unleashed by commercial success are a lot more powerful than the classical models of diffusion suggest. Abernathy and Utterback (1978), Abernathy and Clark (1985), Teece (1986) and Hughes (1994), among others, have rightly underlined that in addition to the rigidities which may affect individual workers or machines and individual intermediaries or users, there are *systemic rigidities* of a much greater scope and importance. Every successful generic technology has a complex web of complementary technologies woven around its core. Once such an integrated and expensive, in terms of purchasing and using, technological system is in place, its momentum becomes enormous. Consequently, once a specific new technology becomes part of a dominant system, it will become increasingly difficult to dislodge, even by more worthy alternatives. This is depicted in Figure 4.2 where technology B (the inferior one) is foreclosed by the entrenchment of technology A (the superior one), which will only be displaced by the much superior technology C.

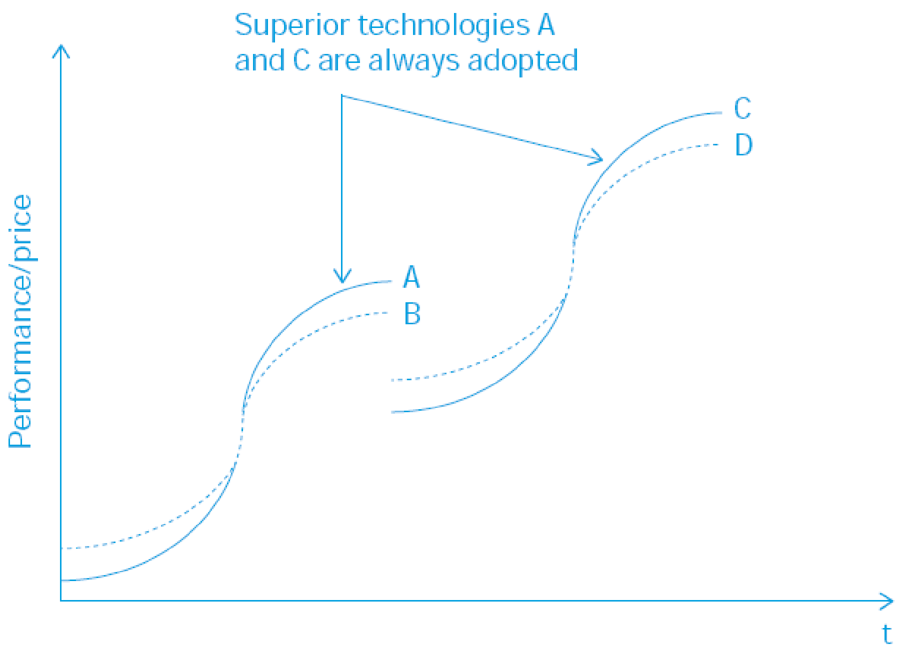

Figure 4.2. Uncertain and Inertial Technological Trajectory: superior technologies A and C are always adopted.

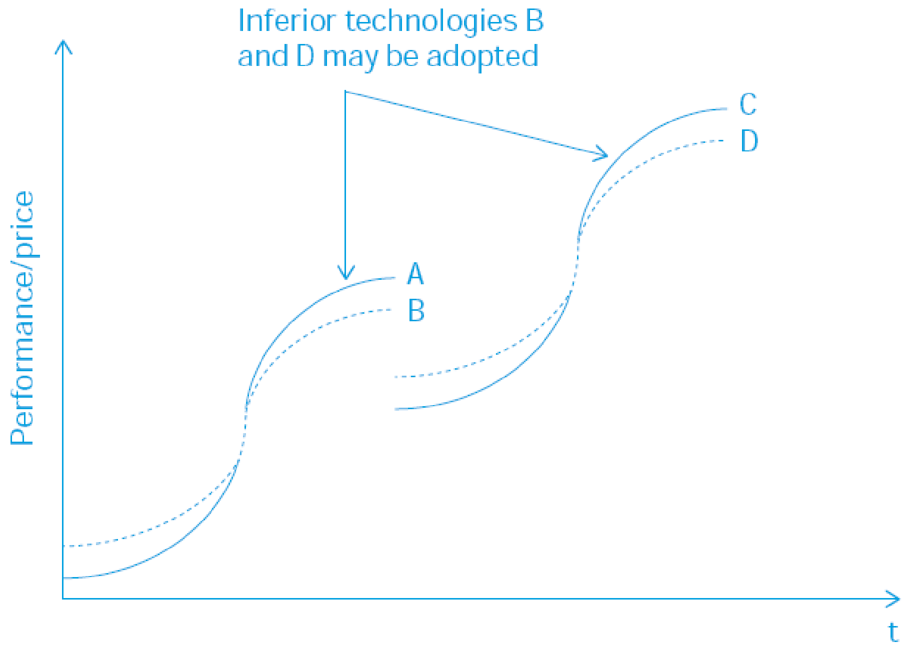

Figure 4.3. Indeterminate Technological Trajectory: Inferior Technologies B and D may be adopted.

Uncertainty and inertia can combine to cause decisive first-mover advantages, which may grant an unassailable market dominance to an early technological trajectory. And yet, these authors do not go far enough in recognizing that in the real world, optimal technological cycles, trajectories and discontinuities, such as options A and C in Figure 4.2, are not unalterable realities. They largely ignore and or minimize the self-reinforcing (and not simply inertial) nature of commercial success and the consequent unpredictability of technological evolution in general and of 'dominant designs' in particular. With this, they also overlook that a technology's success is tributary to the competitive decisions (often arbitrary and myopic) of the major players in an industry, as they are to any set of exogenous technical parameters.

Abernathy and Utterback did not believe in the research lab as an optimal selector of new technologies and they did question the optimality of selection by the market, but only to a small degree. The early articles written together by Abernathy and Utterback are thoroughly ergodic markets-oriented (Abernathy and Utterback, 1978). Nowhere in them is it hinted that a dominant design might not be optimal or that its lasting power might not be inevitable. Later on, Abernathy and Clark (1985) made a strong case for contingency in the maturity and decline of technologies and industries. They thus rejected in no uncertain terms the deterministic view of technological life cycles. Paradoxically, however, their emphasis on historical contingency did not extend to the emergent phase of a new generic technology. They continued to suggest or imply that within a given generic technology, a specific 'dominant design' will be chosen strictly on the basis of its relative merit. As for Utterback, in his more recent writings (Utterback, 1993; Utterback and Suarez, 1994), he fully acknowledges that indeterminacy characterizes both the emergence and the decline of a generic technology.

Our observations on technological competition have shown that markets, in the presence of increasing returns to adoption, tend to become very unstable and tipping –i.e., to discriminate sharply between winners and losers – often on the basis of minimal, perhaps almost random market share differences among the various offers and regardless the relative merit and potential of a new technology. From two comparable competing technologies A and B (see Figure 4.3) in a market characterized by unbounded increased returns to adoption, only one will win the race for dominance (lock-in). But, a priori, it is hard to determine which technology will tip the market (indeterminacy). Furthermore, it is not always sure that the market will select the superior option (sub-optimality). Then, the market is not necessarily an optimal selector of optimal technologies.

Another aspect the work on technological competition under increasing returns has shed light on is related to the implications of market instability for the management of risk.

Ergodic models tacitly assume that new technologies are cheap to develop, hard to improve, stand alone, easy to appraise, easy to use, and strongly protected through patents. This implies technologies arrive at the market full developed before diffusion and that the process of development and the diffusion itself can be separated from each other. Ergodic models, then, introduce a very limited level of uncertainty.

In the models of indeterminate technological change, in contrast, new technologies are expensive to develop, subject to further improvements, often due to systemic nature, difficult to appraise and use and with weak patent protection. Therefore, their adoption become self-reinforcing not only for the *reputation effect* of market success, but also for the significant improvements that technologies accumulate during their spreading. As the *diffusion process confers value* to technologies and not only conspicuousness, a technology that initially did not deserve being chosen may end up meriting it. Consequently, according to Foray (1989), the market becomes not only a selector of adequate technologies, but a creator of dominant, superior ones. Under these circumstances, technological sponsors do not face an information problem. but an indeterminate scenario, as Arthur (1994) has indicated. Thus the most effective way to manage totally contingent and unpredictable results in unstable environments is to invest aggressively in market share as the market takes off.

4.6. Standards and Increasing Returns

Early models of technology diffusion explicitly recognized the general notions of adoption externalities and self-reinforcing dynamics. The epidemic model specified that the diffusion of a new technology, much as that of a potent virus, would be a self-reinforcing process since every user of the respective technology would turn into one of its ardent promoters. What these models, however, failed to recognize was that diffusion can increase considerably the value, and not simply the reputation, of the new technology. In other words, the early models were too focused on the information externalities and neglected to take into account another source of selfreinforcing in diffusion, namely, *increasing returns to adoption.*

In recent years, technological competition and the emergence of a technological monopoly over a whole market has been the privileged topic of the literature on standards (David and Greenstein,1990). This literature has identified three processes by which technological standardization can be attained: (a) government regulation through mandated standards, (b) voluntary agreements through formal or standardization committees, and (c) market competition. The first two processes, often called *de jure* standardization, usually result in a standard with public-good characteristics. Standards selected through the market –*de facto* standards– on the other hand, are usually owned by a firm, which can therefore exclude other firms from its use. Every standardization process has its own theoretical interest, but here we will focus exclusively on *de facto* standardization, which is identified with the economic literature on technological competition under increasing returns.

4.7. NETWORK EXTERNALITIES

At the basis of what we know about technological competition is the literature on network externalities (Katz and Shapiro, 1992,1994; Economides, 1996) in which market size relates to increasing returns, and benefits grow with the size of competing networks.

Just as scale economies, learning, or reputation effects, positive network externalities are a self-reinforcing diffusion dynamic. Network effects, however, differ from the other self-reinforcing mechanisms in several important respects: First, while the benefits of scale economies, learning and (some) reputation effects can only be reaped at the time of purchasing the product in question, most of the benefits accruing from network externalities can be enjoyed well past the point of purchase and throughout the entire life cycle of the product. Second, network effects are considerably forward looking and less bounded and therefore more powerful than scale and learning effects. In fact, because they cast a shadow into the future, network effects can combine with reputation effects to create extremely powerful and lasting self-reinforcing dynamics in market success. Since most of the benefits accruing from network externalities can be enjoyed throughout the full life cycle of a product, new users faced with a multiplicity of competing technical options will have to make their choices not simply in function of what the majority of post purchasers have already chosen, but also in function of what the majority of *inter-face* users are likely to choose. Interestingly, while very pessimistic user expectations about an overall emerging market can be self-dampening, optimistic expectations about the success of a specific technical format in a battle of standards could easily become self-fulfilling. The greater the number of people who think that a given technical option is likely to become dominant, the more new users will side with it and the more likely and pronounced the dominance will in fact become. Third, while scale economies and learning can only be a source of increasing returns to adoption and while users' learning costs (or switching costs) can exclusively be a source of inertia, both reputation effects and network externalities, in contrast, can act as both strong inertial and increasing returns to adoption.

4.8. REFLECTIONS ON INCREASING RETURNS IN INDUSTRIAL COMPETITION

The analysis so far done has gradually recognized the central role of increasing returns mechanisms in generating and sustaining dominant firms and technologies. Clearly, while scale economies, the resource-based loop, the Schumpeterian (innovation-based) loop, reputation, and the different categories of learning help to explain some of the most basic occurrences of dominant firms and slanted industrial structures, only increasing returns to adoption – a notion intrinsically connected to Schumpeterian economics - can explain most of the instances of technological dominance which we see in contemporary high-technology markets. Thus the integration of all these increasing returns mechanisms, as Figure 4.1 shows, results in a quite complete explanation of industrial competition.

Such an explanation combines the self-reinforcing loops based upon resources and innovations (loops A and B), scale economies, learning, and reputation (loops C) with the loops based upon increasing returns to adoption. These loops are of two kinds. A set of further loops, composed of a mesh of scale economies, learning, reputation effects, infrastructure effects, and network effects, links increasing competitive advantages with increasing returns to market share. A last set of loops indicates that, if increasing returns to adoption are present and considerable, market share becomes a strategic asset well worth investing on in an aggressive manner through vigorous production capacity expansion, price reductions, infrastructure development, and alliances with manufacturer of complementary technologies, products, and services.

As a model of increasing returns mechanism in industrial dynamics, the one described here seems to be rather comprehensive. It can explain the polarized outcomes that are common in most industrial sector and describe business competition as a dynamic and cumulative process. A final, but not less interesting, feature of this model of industrial competition is that it is general, in the sense that it is capable to describe simultaneously the three levels of industrial competition. It elucidates how technological adoption, the number and growth of firms, and industrial structure combine and cause each other. In other words, the general model of industrial competition gives a picture of how industrial competition is a process in which technological competition affects the size of the firms competition in a given industry and how the growth of the firms, in turn, influences the structure of that industry. That is why technological adoption goes between conduct and performance in the chain of causality that leads from the size of the firms to industrial structure.

4.9. Increasing Returns, Structural Change and Development Paths: An Example

The seminal article formalizing the ‘big push’ theory of industrialization is that of Murphy, Shleifer and Vishney (1989). In their model, firms choose between a constant returns and an increasing returns of technology based on their expectations of demand. However, these choices spill over into aggregate demand creating a strategic interaction among sectors in their technology adoption decisions. Thus, under certain conditions, there exist two equilibria: with all firms choosing the constant returns or all choosing the increasing returns technology. Clearly, in the latter equilibrium, all households are better off.

While the Murphy, Shleifer and Vishny model shows how increasing returns (and a wage effect) aggregate to strategic complementarity among sectors, it does not lend itself readily to the debate concerning the degree of balance in industrialization policy.

First, the static content leaves open the question of whether the intervention should take the form of anything more than indicative planning. Second, the most commonly discussed policy instrument in the industrialization debate is the subsidization of investments.

However, in the Murphy, Shleifer and Vishny example, use of this instrument biases one toward a more unbalanced policy. To see this, observe that it is the role of the government to facilitate a move to the industrializing equilibrium. This means that the government must subsidize a sufficient amount of investment to make it profitable for all sectors to adopt the modern technology. Given the binary choice set, there then exists some minimum critical mass of sectors that must be targeted to achieve a successful transition. A greater range of successful industrialization policies might be more plausible, however, if firms had the choice of a wider variety of technology to choose from. One might suppose that targeting a large number of sectors to modernize a little and targeting a small number of sectors for more radical modernization might both generate a big push. Thus, to consider the balanced approach properly, a greater technological choice space is required.

What would be the choice variables available to the government provided it would be able to pick up what is likely to be increasing returns industries in the future?

First, in each period, the government can choose the set of firms that it targets for structural change. Second, for each targeted firm, the government can choose a target level for ‘increasing returns industry’ modernization in the period. Along this vein, the government could choose to target the same number of firms in each period but induce those firms to modernize gradually over time. Or in contrast, the government chooses a single level of modernization to occur across all firms and all periods. It then targets a mass of firms each period for entry and modernization. This means that industrialization policy is solely characterized by the critical mass of sectors targeted, and the target level of modernization.

Given a parametrized development path, the most significant parameter represents the strength of increasing returns in the technology adopted by industrial sectors which generates a rationale for ‘big push’ intervention.

A ‘big push’ can be activated if the economy is stuck in a ‘development trap’ from which an escape could be made through sufficient coordination of decisions by input producers. For a developing economy in its early phase a ‘poverty trap’ is a special case of a ‘development trap’ defined by Barro and Sala-i-Martin (1995, 49) “…as a stable steady-state with low levels of per capita output and capital stock.” This is a trap because, if agents attempt to break out of it, the economy has a tendency to return to the low-level steady-state. Only by a very large change in their behavior can the economy break out of the poverty trap and move to the high(er) income steady-state.

To evaluate the economic characteristics, i.e., the strengths of complementarities and increasing returns, should affect the government’s policy choices and industrial policies (Gans, 1994; Wydick, 2008, Chap.3).

Big Push theories of industrialization could lead to ‘development traps’ if sequential industrialization would add more diminishing returns than increasing returns industries which could be a result of government’s coordination failure. When a development trap is purely the result of coordination failure, to escape from the trap, would require the government to synchronize the expectations of individual agents with targeting investment in industrialization activities. If a government were to announce that firms should modernize to a certain degree, even if this were believed perfectly by individuals and firms, each firm might still have an incentive to wait before investing. In that case, the optimistic expectations by the government would not be realized and the policy would be ineffective. Irreversibility and the time lag of production mean that history rather than expectations matter (Krugman, 1991). The previous level of industrialization determines what path the economy will take in the future. This is why it is difficult to characterize the industrializing paths of the economy.

4.10. INDUSTRIALIZATION POLICIES AND DEVELOPMENT

In the context of a big push development strategy the government faces a tradeoff between the number of sectors it targets and the degree to which it wishes them to modernize, that is, it chooses the critical mass of sectors that must be targeted at any point in time in order to generate an escape from a development trap and to achieve increasing returns. Let's take a simple case where the industrialization policy takes the form of a 'big bang,' that is intervention occurs for one period only granting that the resources exist in that period to allow for such a policy. This means that the industrialization policy is solely characterized by the critical mass of sectors targeted s^* and the target level of modernization f.

Suppose naturally that individual transition costs are non-decreasing in f, the optimal critical mass in terms of f can be described by the path

$s^*(f,\phi) = (f + 1)^{\theta(1-\sigma)} [((1-\delta)(\sigma-1)/ L\ \bar{\lambda})^{(1-\sigma)/(\sigma-1-\sigma)} - s_I)] + s_I$ with s_I as the basic input varieties of the industrial economy.

Substituting this into the objective function with cost c (f,1; s_I, ϕ), the 'big bang' industrialization policy problem becomes

$$\min_f (f + 1)^{\theta(1-\sigma)} [((1-\delta)(\sigma-1)/ L\ \bar{\lambda})^{(1-\sigma)/(\sigma-1-\sigma)} - s_I)]\ c\ (f,1; s_I, \phi)$$

where use is made of the symmetry of the cost functions and the fact that $s^* - \underline{s}_I$ firms are targeted. ϕ could represent any given exogeneous parameter, i.e., σ, θ, δ, α, or $\bar{L}$, λ a given parameter linked to $\bar{L}$.

In designing an optimal industrialization policy it shows that a cost minimizing policy in the industry transition entails setting certain development model (exogeneous) parameters such as labor productivity improvements (θ), upstream firms discount future earnings (δ) the fixed size of the labor force ($\bar{L}$), the number of basic industrial sector varieties (s_I), the product linkages between intermediate input producers (σ), and the use of the intermediate input composite (α), the latter two showing a certain degree of interaction referring to as the returns to specialization (Romer, 1986). Discussing these parameters qualitatively in terms of comparative statics would indicate industrial change. Raising any of these parameters θ, δ, $\bar{L}$, and s_I increases the marginal returns to upstream firms in both their entry and modernization decisions.

Raising θ means that sunk costs are translated into labor improvements more effectively. Similarly, since the costs of modernization and entry are carried today and most of the returns occur in the future, the more likely they are to undertake those actions. A large market, a higher $\bar{L}$, also raises the marginal return to entry and modernization. Finally, more industrial varieties mean that the past level of industrialization is greater, thereby, reducing the marginal costs of inducing firms to adopt more modern technologies. Given this, the responsiveness of firms to inducements by the government is enhanced when any of these parameters is raised.

Therefore, the higher are these parameters, the fewer firms need to be targeted to facilitate an escape (from a development trap).

Of these parameters θ has probably received the most discussion. In many ways, this parameter represents the strength of increasing returns in the technology adopted by upstream producers. This is because higher levels of θ imply that, when they choose to modernize, upstream firms will choose technologies involving greater sunk (or fixed) costs.

Therefore, the higher are these parameters, the fewer firms need to be targeted to facilitate an escape (from a development trap).

Of these parameters θ has probably received the most discussion. In many ways, this parameter represents the strength of increasing returns in the technology adopted by upstream producers. This is because higher levels of θ imply that, when they choose to modernize, upstream firms will choose technologies involving greater sunk (or fixed) costs. Therefore, while one requires some degree of increasing returns or economies of scale in production to generate a rationale for a 'big bang' intervention, the stronger are those increasing returns supports a more unbalanced industrialization policy. This relates back to arguments made on balanced vs. unbalanced growth in development policy (Easterly, 2002).

Of the three other parameters, only the discount rate δ seems to have been given a potential role in the past debate on industrialization policy. Matsuyama (1992) interprets the discount rate as a measure of the effectiveness of entrepreneurship in coordinating investment, with a low discount rate indicating the existence of greater entrepreneurial resources. If so, then the above result seems to imply that with a relative scarcity of entrepreneurial talent a more balanced approach should be followed.

The comparative statics results for α and σ require more restrictions because each of these has two effects. On the one hand, lowering σ and increasing α raises the strength of strategic complementarities among upstream sectors. This tends to favor a more balanced growth approach. On the other hand, α and σ each affect the marginal returns to entry and modernization of firms. The second effect reinforces the first and leads to a more balanced strategy, that is, lowering σ and lifting α increase the marginal returns to entry and modernization. A lower σ also implies stronger technical complementarities. This effect is sometimes referred to as the returns to specialization (Romer, 1987). The consequence is that a lower σ raises the marginal returns to employing a greater variety of inputs in production. The higher is σ the weaker are the linkages among intermediate input sectors. Conversely, stronger linkages between sectors raise the marginal return to targeting an additional sector for change supporting the arguments of the balanced growth strategy.

Looking at α, it is a measure of the appropriability of the returns from supply an additional intermediate input. As Romer (1994) discusses, the larger is α, the greater is the surplus gained by intermediate input producers from the employment of their product in final good production. Therefore, producers of inputs targeted in an industrialization policy are more likely to react positively (in terms of adopting better technology) when the appropriable returns from the introduction of their variety is larger. This effect would tend to favor a more unbalanced approach as α increases.

Summarizing, we have outlined the role of several parameters in influencing the kind and degree of balance in industrialization policy. Factors addressed in the earlier literature such as strength of linkages, increasing returns and entrepreneurial resources all influence the composition of the 'big push.' By considering a 'big bang' policy, some results are possible. For instance, as developed at length in this chapter, strong increasing returns in conjunction with weak sector linkages tend to favor a more unbalanced approach in order to minimize costs.

4.11. INCREASING RETURNS MECHANISMS AND TECHNOLOGICAL COMPETITION

The strategic importance of increasing returns to technology adoption is unquestionable. In a strictly ergodic market technological options eventually obtain the share of the market they deserve in proportion to their value and technical possibilities. In non-ergodic situations, in a tipping market, on the other hand, the winner takes all or most and the losers (no matter how worthy and how many of them there are) loses all or much. Because of this, the introduction of factors causing tipping markets determines the outcomes of technological competition. This framework not only captures the interplay of institutional arrangements, resources and network of firms and industries in industrial competition, but also delineates very concrete regularities, which can provide us with a simple but powerful tool to explain endogeneously, and in a dynamic way the firms' growth, industrial structure, and technological competition. This chapter suggests a Schumpeterian model of industrial competition and growth extending through increasing returns mechanisms to neoclassical models of industrialization. In contrast to the approaches that underline a specific aspect of industrial competition and/or base their explanation on a reduced number of factors and against conventional economics that overlooks increasing returns mechanisms, this model links the increasing returns mechanisms that determine endogenously interfirm asymmetries and the kind of industrial structure which emerges during this competition process. This framework also emphasizes the fact that the emergence of dominant firms and the evolution of industrial structure are strongly intertwined with the process of technological change and diffusion. One of the most important contributions of the work on increasing returns is in having shown that the emergence and persistence of technological monopolies is not an exogenous datum, largely determined by scientific and technical parameters, but is strongly influenced by vibrant market forces stemming from self-reinforcing mechanisms. In the presence of strong, global and long lasting increasing returns, the actions and omissions of the main actors in the industry in question affect considerably the final result of technological competition. To the extent to which these actors are capable of fully perceiving and exploiting strong increasing returns in emerging markets. they can ensure the entrenchment of their technology as the industrial standard by investing in those strategies that bring about a growing market share. Once entrenched, and to the extent they are capable to exploit inertial forces, established firms can ensure the persistence of their technologies well beyond the time warranted by their relative technical value.

We also show that technological monopoly should be attributed to increasing returns in general, but it is network externalities in particular what has caused some important markets to be dominated by a technological monopoly. In fact, at the level of technological competition, the only thing the other sources of increasing returns to adoption do is to exacerbate the implications of network effects, but they do not turn, isolated from network effects, a market tipping. The most obvious and direct reason for technological monopoly is that the components of a given network are compatible and constantly interconnected. The telephone and the fax are examples of networks where physical interconnection and compatibility have led to technological monopoly.

Winner-takes-all markets are associated with cases where there is (often intense) competition in innovative activity but the future market is such that competition in it is, over a reasonable timeframe, not sustainable. Thus firms compete to attain a position of dominance.

Perhaps the most famous example of a winner-takes-all market is that for operating systems on desktop PCs or other devices. It is instructive to recognize that this market benefits from massive economies of scale in production protected by intellectual property rights (very low marginal cost of supply compared with very large fixed costs of initial product development) and substantial economies of scale in consumption (due in large part to the network effects associated with the relationships between the operating system market and the related applications software markets).

Identification of such markets is important because it affects the focus of competitive concerns.

Most obviously, if there are strong grounds to believe that a future market is a winner-takes-all market, it is perhaps not appropriate for a competition authority to block a merger or agreement between firms on the basis that this will create a dominant position or lessen competition in this future market. By definition, the nature of the market is such that its existence guarantees that a firm will be dominant on it, at least in the medium term. (This illustrates an important point relevant to wider issues in competition policy: it is typically better to have a situation where a firm is dominant in a relevant market than for that market not to exist at all.) Instead, any intervention must be based on the premise that the merger (agreement) lessens or distorts competition on some other, perhaps related market, or in competition in the innovative activity associated with the winner-takes-all market.

Similarly in dominance cases, if we anticipate that a market is subject to winner-takes-all properties, then it is difficult to establish a case that a firm has abused its dominant position in monopolizing this market - the market is naturally prone to monopolization. Rather, analysis of an alleged abuse of dominance associated with this market should focus on how a dominant position in a related market (perhaps an access market) could be used to distort competition in the innovative activity associated with the winner-takes-all market, or how a dominant position in the winner-takes-all market could be abused to maintain that position - in effect used to distort competition in the innovative activity associated with the future generation of that market.

Indirect network effects may also tilt the market in favor of one of the competing technologies. In the videofilm industry, for instance, because of strong, long lasting, and global network externalities, technological monopoly emerged and a product cycle has lasted about 20 years. The strength, duration and scope of increasing returns in the VCR markets are the direct consequence of particular technical characteristics of the competing VCR technologies. These technologies were quite similar and mature, the usage cycle of their compatible content -videofilms– is very short, and their potential substitutes have not had large enough advantages so as to replace them. In the video game industry, in contrast, because of weak, short lasting, and local network externalities due to the incomparability of videogame technologies, long usage cycle of videogames, and the proliferation of new, more powerful new formats, different technologies have tended to share the market according to their intrinsic value (Caves, 2000, Chapter 13).

A comparison between the videogame industry and the videofilm industry, then allows us to show that compatibility constraints are not a sufficient condition for a virtual network technology to become a monopoly: it is necessary, besides, that usage life of content is short and core technologies are incomparable. But when the usage cycle of content is long and core technologies are comparable, any virtual network technology becomes less systemic. As the strength, duration, and scope of network externalities are reduced considerably because of a long-lasting usage cycle and technology comparability, competing virtual network technologies end up sharing the market in proportion to their value.

Comparative assessments of this sort advance our understanding of the technical and institutional factors which are likely to affect the nature of technological competition. In doing so, we add to the dimensions of *strength* and *duration*, the dimension of *scope* of increasing returns. The distinction between strength, duration and scope is useful to realize that, contrary to popular and academic literature, a market for virtual network technologies with content dimensions not necessarily will end up with a technological monopoly. With this distinction in mind and against those who think that strong indirect network externalities always act as tipping mechanisms, we also can show that strong indirect network externalities are compatible with fairly ergodic market dynamics, if the scope of such externalities happens to be rather narrow.

By taking the telecommunications or the videofilm industry, for example, we can show that strong network externalities are *necessary* but not *sufficient* conditions to produce technological monopoly. Short usage life of content and technology incomparability are technical and necessary conditions for technological monopoly in software intensive virtual network technologies to happen. But these technical aspects of virtual network technologies are not a sufficient condition to produce technological monopolies. In both Chapters 2 and 4 we argued that network externalities require not only to have high levels of strength. but also to be global in scope. Under certain institutional conditions strong indirect network externalities may be rather localized, which leads to very ergodic market results. In these conditions markets are shared by the competing technologies according to their intrinsic value.

Technical and institutional factors causing different levels of strength, duration and scope of increasing returns to adoption are relevant to determine whether a market is tipping or ergodic. This has some implications. A first implication is that not all network technology is equally systemic. If there are strong network externalities but with a local scope, the systemic nature of a network becomes rather limited. In this case the systemic nature of the VCR network would be rather local compared with the actual network, which is global in nature.

Strength, duration, and scope of increasing returns are also useful to determine in a more detailed way the nature of *cooperation.* When network externalities are strong and global, content-intensive virtual network technologies become rather systemic. In these circumstances, the main sponsors of the competing technologies may produce some components of the system but the rest of it may be out of their reach. For instance, a PC producer may be incapable to produce software or microprocessors; and microprocessor producers may not be able to produce software or hardware. VCR producers cannot produce films, in the same way as film producers cannot produce VCRs. Consequently, technological competition in markets characterized by strong and global increasing returns is more in connection with complex networks of firms than with conventional industrial array of firms producing homogeneous products. In contrast, in markets with weak and local network externalities, competition takes place mainly between firms than between networks of firms. This is so, because in this kind of markets products are not systemic.

REFERENCES

Abernathy, W. J. & Utterback, J. M. (1978). "Patterns of Innovation in Technology," *Technology Review*, *10*, 40-47.

Abernathy. W. J. & Clark, K. B. (1985). "Innovation: Mapping the Wind of Creative Destruction," *Research Policy*, *14*, 3-22.

Aghion, E. & Howitt, E. (1998). Endogeneous Growth Theory, Cambridge, Ma.: MIT Press.

Arrow, Kenneth (1962). "The Economic Implications of Learning by Doing," *Review of Economic Studies*, *29*, 155-173.

Arthur W. B., Ermoliev, Y. M. & Kaniovski, I. M. (1987). "Path-dependent Processes and the Emergence of Macro-structure," *European Journal of Operational Research*, *30*, 294-303, also in Arthur (1994), 33-48.

Arthur, W. B. (1989). "Competing Technologies, Increasing Returns and Lock-in by Historical Events," *The Economic Journal*. *99*, 116-131, also in Arthur (1994), 13-32.

Arthur, W. B. (1990a). "Positive Feedbacks in the Economy," *Scientific American, February*, *22*, 92-99, also in Arthur (1994),1-12.

Arthur, W. B. (1990b). "'Silicon Valley' Location Cluster. When Do Increasing Returns Imply Monopoly," Mathematical Social Science 19, 235-51, also in Arthur (1994), 49-67.

Arthur, W. B. and D. A. Lane, 1993, "Information Contagion," Structural Change and Economic Dynamics 4, 81-104, also in Arthur (1994), 69-97.

Arthur. W. B. (1994). Increasing Returns and Path Dependence in the Economy. *Ann Arbor*, The Univ. of Michigan Press.

Arthur, W. B. (1999). "Complexity and the Economy," *Science*, *284*, 107-109.

Barro, R. A. & Sala-i-Martin, X. (1995). Economic Growth, Cambridge, Ma.:MIT Press.

Baumol, W. J. (1993). Entrepreneurship, Management, and the Structure of Payoffs, Cambridge, Ma.: MIT Press.

Caminal, R. & Vives, X. (1996). "Why Market Shares Matter: an information-based theory," *RAND Journal of Economics*, *27*(2), 221-239.

Caves, R. E. (2000). Creative Industries, Contracts between Art and Commerce, Cambridge, Ma.: Harvard Univ. Press.

Chandler, A. D. (1990), Scale and Scope: Dynamics of Industrial Capitalism., Cambridge, Ma.: The Belknap Press of Harvard University Press.

Clark, K. B. & Fujimoto, T. (1991). Product Development and Performance, Strategy, Organization and Management in the World Auto Industry, Boston: Harvard Business School Press.

Cohen, W. M. & Levin, R. C. (1989). "Empirical Studies of Innovation and Market Structure," in Schmalensee, R. and R. D. Willig. eds., Handbook of Industrial Organization. Amsterdam: North-Holland.

David, P. & Greenstein, S. (1990), "The Economics of Compatibility Standards: an Introduction to Recent Research," *Economics of Innovation and New Technology*, *1*, 3-41.

Easterly, W. (2002). The Elusive Quest for Growth, Cambridge, Ma.: MIT Press.

Economides. N. (1996). "The Economics of Networks." *International Journal of Industrial Organization*, *25*, 1-18.

Feller, W. (1957). An Introduction to Probability Theory and its Applications, Vol. I, New York: John Wiley & Sons.

Foray, D. (1993)."Standardisation et Concurrence: Des Relations Ambivalentes" *Rev. d'Economie Industrielle*, *63*, 30-45.

Francois, P. & Lloyd-Ellis, H. L. (2002). "Animal Spirits through Creative Destruction," Disc. Paper, Economics Dept., Queen's Univ. Kingston, Ont., Canada, May, mimeo.

Freeman, C. (2003). "A Schumpeterian Renaissance?" Science and Technology Policy Research Unit, Univ. of Sussex, Brighton, SEWPS No.102, July.

Gans, J. (1998). "Industrialization with a Menu of Technologies: Appropriate technologies and the 'big push,' " *Structural Change and Economic Dynamics*, *9*, 333-348.

Gottinger, H. W. (2003). Economies of Network Industries, London: Routledge.

Henderson, B. D. (1975). "The Market Share Paradox," in Y. Brozen, ed. The Competitive Economy, Morristown, N. J.: General Learning Press, 286-287.

Hirsch, W. Z (1956). "Firm Progress Ratios," *Econometrica*, *24*, 136-143.

Katz, M. & Shapiro, C. (1992). "Product Introduction with Network Externalities," *The Journal of Industrial Economics*, *40*, 55-84.

Katz, M. & Shapiro, C. (1994). "System Competition and Network Effects," *Journal of Economic Perspectives*, *8*, 93-115.

Krugman, P. R. (1991). "History versus Expectations," *Quarterly Journal of Economics*, *106*(2), 651-667.

Matsuyama, K. (1992). "The Market Size, Entrepreneurship, and the Big Push," *Journal of the Japanese and International Economies*, *6*, 347-364.

McAfee, R. P. (2002). Competitive Solutions, Princeton: Princeton Univ. Press.

Murphy, K. M., Shleifer, A. & Vishny, R. W. (1989). "Industrialization and the Big Push," *Journal of Political Economy*, *97*, 1003-1026.

Nelson, R. & Winter, S. G. (1982). An Evolutionary Theory of Economic Change, Cambridge, Ma.: Harvard Univ. Press.

Nelson, R. (1986). "Evolutionary Modeling of Economic Change," in Stiglitz, J. and G. F. Matthewson, eds., New Developments in the Analysis of Market Structure, Cambridge, Ma.: MIT Press.

Porter, M. E. (1980). Competitive Strategy, New York: Free Press.

Radner, R. (1986). "The Internal Economy of Large Firms," Annual Conference of the Royal Economic Society, *Supplement to the Economic Journal*, *96*, 1-22.

Roberts, J. (2004). The Modern Firm, Oxford: Oxford Univ. Press.

Romer, P. M. (1986). "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, *94*, 1002-1037.

Romer, P. M. (1987). "Growth Based on Increasing Returns due to Specialization," *American Economic Review*, *77* (2), 56-62.

Romer, P. M. (1994). "New Goods, Old Theory, and the Welfare Costs of Trade Restrictions," *Journal of Development Economics*, *43*, 5-38.

Rosenberg, N. (1982). Inside the Black Box: Technology and Economics, Cambridge: Cambridge Univ. Press.

Schumpeter. J. A. (1942). Capitalism. Socialism. and Democracy, New York: Harper.

Shapiro, C. & Varian, H. R. (1999). Information Rules, Boston: Harvard Business School Press.

Teece, D. (1986). “Profiting from Technology Innovations,” *Research Policy*, *15*(6), 285-305.

Tirole, J. (1988). Theory of Industrial Organization, Cambridge, Ma.: MIT Press.

Utterback, J. M. (1994). Mastering the Dynamics of Innovation, Boston: Harvard Business School Press.

Utterback, J. M. & Suarez, F. F. (1993). “Innovation, Competition and Industry Structure,” *Research Policy*, *22*, 1-21.

Wydick, B. (2008). Games in Economic Development, Cambridge: Cambridge Univ. Press.

Young, A. A. (1928). “Increasing Returns and Economic Progress,” *Economic Journal*, *38*, 527-542.

Chapter 5

NEW PERSPECTIVES ON INDUSTRIAL AND ECONOMIC GROWTH: A REVIEW

5.1. PREVIEW

Technological innovation has an important role to play both as part of the competitive process and as a key driver of networks, industrial growth and development. However, perhaps surprisingly, for a long time innovation was considered a field where economists, as opposed to other scientists and engineers, only viewed this as a 'black box' (Landau and Rosenberg, 1986). The forces underlying the process of technological change were believed to be substantially independent of economic incentives and mostly affected by the exogenous evolution of scientific knowledge and its technological applications.

It has been only recently, mainly over the last three decades, that the economic forces behind technological innovation have started to be investigated in more detail (Barro and Sala-i-Martin, 1995). As a result, a wide variety of theories and models, sometimes very diverse in spirit, describing the economics of innovation are now available. All these theories share the common aim of providing a conceptual foundation for understanding how innovation affects the economy, how economic forces affect the emergence of technological changes, and the decision-making processes through which technological innovation occurs.

In what follows we provide a rationale for innovations and networks as it relates to industrial growth embedded in a review of the micro-foundations of economic growth.

Generally, traditional models of innovation focus on the study of firms' incentives to invest resources in Research and Development (R&D) activities. Game-theoretical models (GTMs) as developed in the industrial organization literature have investigated firms' R&D decisions in strategic environments (Tirole, 1988). Endogenous growth models, to be discussed, have developed the study of market dynamics in models that explain the relationship between firms' investments, innovation and economic growth.

GTMs suggest that there are two main forces that underlie firms' investment in R&D: the search for higher profits and the threat posed by falling behind potential innovating rivals. They study these forces in a variety of market situations and address issues such as the interplay between innovation and market structure, the dynamics of competition and the nature of the relationship between intensity of competition and innovation.

Despite the absence of general results, for universal application in a policy setting, these models are certainly useful tools to understand firms' incentives to invest in R&D activities in strategic environments and to suggest what main factors may be central in shaping the nature of dynamic competition. For instance, these models suggest that in order to understand R&D investments in strategic environments it is necessary to understand how innovation may affect profits both of successful and non-successful innovators. One perspective captures the idea that firms want to innovate to increase their profits; another captures the idea that firms want to innovate to maintain competitiveness.

Dynamic competition may be characterized by persistent dominance of the incumbent leader or by action-reaction whereby incumbents are overtaken by a rival whose incumbency itself may only be transitory. The nature of market dynamics depends on a number of factors, such as the type of innovation, i.e., radical or non-radical, the uncertainties involved in R&D activities, the nature of patent protection and of knowledge spillovers, the intensity of product market competition and other strategic factors (Gottinger, 2006).

When the relationship between competition and innovation is investigated, it is necessary to be clear what the notion of intensity of competition describes and how this relates to market structure. Indeed a market where competition is tougher may be more concentrated simply because inefficient firms cannot survive. There may be a trade-off between the intensity of static competition and innovation. In general, the relationship between intensity of competition and innovation need not be monotonic at all. Endogenous growth models (EGMs) have been developed along the earlier game theoretical literature on innovation in the context of studies that seek to explain the relationship between innovation and economic growth (Reinganum, 1989). These models suggest that innovation, resulting from intentional R&D investments by profit-maximizing firms or simply by unintentional learning-by-doing, is a fundamental driver of economic growth in the long run.

Early EGMs stressed the importance of ex-post rents for innovation: competition would have a detrimental effect on innovation by decreasing the rents that an innovator would be able to appropriate. More recent models have emphasized another mechanism by which competition affects innovation: tougher competition may increase the incentives of firms to innovate in order to escape from fierce competition (Aghion and Griffith, 2005). These recent studies suggest that the relationship between competition and innovation may not well be monotonic and that instead, one should expect an inverse relationship: when competition is low, an increase in competition would foster innovation; the reverse would happen when competition is fierce.

The result that competition may be conducive to innovation is also obtained in studies where the traditional behavioral assumption of profit-maximizing firms is relaxed. When principal-agent considerations are introduced to explain managers' behaviors, another mechanism by which competition may favor innovation is suggested: the speed of innovation may be retarded by the slack of managers who tend to avoid private costs associated with innovation. When competition intensifies, the higher threat of bankruptcy may force managers to speed up the process at which new ideas are adopted. Hence, competition may be conducive to faster rates of innovation.

At an aggregate level, the evolutionary approach to the study of economic growth draws attention to the importance of institutions in the process of economic growth. The key findings of this literature are as follows:

Innovation and its diffusion take place within systems of interconnected organizations and institutions. Important constituent elements of such systems are organizations such as firms, governments and universities. Institutions, with which these organizations interact, reflect laws and statutes (e.g., the institution of patent protection) as well as more abstract elements (e.g., cultural aspects of the economy such as the spirit of entrepreneurial activity). Innovation within the system will depend not only on single institutions but also on the nature and intensity of interactions between the various elements of the system. When considering the interaction between competition and innovation, we must remain aware that the effects of such interaction will depend on the evolving institutional background against which agents in the economy operate.

The existing body of cliometrics and institutional economics draws attention to the following factors relevant to ‘innovation systems’: Economic growth is a result of a complex interaction between institutions and markets. Examples of institutions that are frequently important for innovation are the protection of property rights (for example, against piracy) and patent policy. Innovation plays a critical role in economic progress and development. This involves not only technological innovations but also changes in institutions. Thus, organizations themselves evolve according to changing circumstances. These two last points complement the concepts of systems of innovation found in the macroeconomics of the evolutionary approach.

5.2. Economics of Innovation

By innovation in neoclassical theories, production opportunities are described by specifying a production function, which describes the technological relationship between inputs and output. Technological progress is captured by the ability of a firm, or the economy, to produce, for each combination of inputs used, more output. However, such progress was for a long time essentially considered exogenous, as if it were knowledge that falls ‘unexplained’ from heaven. One may follow Kirzner (1973) that in a market economy knowledge would be provided by ‘entrepreneurship’ and in Schumpeter’s terms the entrepreneur is the innovator (Baumol, 1993, Chapter 1) and he would thrive (only) in a ‘Free-Market Innovation Machine’ (Baumol, 2002) enabled through the capitalist process which, ironically, as Schumpeter predicted at the very end might break down. One can very well see that a socialist system like Russia (during its time as the Soviet Union) has had its significant share of technological inventors but only a negligible share of innovators throughout its history because of a lack of a culture and incentives of entrepreneurship (Graham, 2013). One can identify through the historical biography of Isaacson (2014) where innovators and innovations in major high technology industries are placed.

Neoclassical economists were conscious that the consideration of innovation as an economic phenomenon posed some challenges to established theories of perfect competition and that a trade-off may exist between static and dynamic efficiency. Innovation is essentially the production and commercial application of new knowledge (Arrow, 1962).

Innovation and Dynamic Competition

In the last three decades, the analysis of innovation and technological change has attracted much attention within modern economics, due to the development of two closely related disciplines: game-theoretical industrial organization and endogenous growth theories (Barro and Sala-i-Martin, 1995, I.1-3). As a consequence, many of the themes that were once a feature of the less conventional theories of technological advance have now found their way into economics. GTMs of innovation can be categorized as belonging to one of three broad studies of dynamic competition within new developments in economics that have developed once inherently static neoclassical economic theories into modern analyses of innovation and dynamic competition (Fudenberg and Tirole, 1996). Nonetheless, a strong relationship with neoclassical economics remains in the general methodological approach adopted in these theories, which are characterized by the analysis of equilibrium models where fully rational economic agents make optimizing choices. Such models of dynamic competition are highly stylized and this is also reflected in the definition of innovation adopted. In these models, innovation is usually defined as the reduction of production costs or as the commercialization of new and/or better products. Apparently this definition identifies innovation with technological change and neglects the importance of other forms of innovation, such as managerial, organizational and design changes.

However, notwithstanding the emphasis placed by these models on innovation as the commercial use of new or improved products and processes, this stylized definition of innovation can also be considered to encompass non-technological changes.

Similar to the definition of innovation, also the treatment of the process by which new knowledge is generated is stylized in these models. Knowledge is usually considered as deriving from learning-by-doing or from intentional R&D investments, optimally chosen by profit-seeking agents. The R&D process is captured by the definition of a deterministic or stochastic knowledge production function which related the input used in R&D — notably capital, labor, available knowledge – to the new knowledge produced and/or the timing of innovation as in Baumol (1993, Chapter 7).

GTMs study firms' incentives to invest resources in R&D activities in strategic settings and address issues such as the interplay between innovation and market structure, the dynamics of competition and the nature of the relationship between intensity of competition and innovation.

In general, these models are structured on the fundamental distinction between static and dynamic competition. Static or product-market competition refers to the strategic interaction that arises between firms, taking as given the capabilities of these firms in terms of the variety, quality and production costs of the products in their portfolios. Dynamic competition refers instead to the strategic interaction among rival firms to change these capabilities for their advantage. This distinction is usually captured by the structure of game-theoretical models: firms are considered to play a game where they first invest in innovative activities and then compete at the product-market level. The distinction is a fictitious one to a certain extent, since in reality both dimensions of competition are continuously overlapping, but it is a convincing and powerful conceptual categorization since it allows distinguishing activities that occur over different time horizons, i.e., short-term price competition versus long-term dynamic competition. When firms compete in the short-term they have to take as given a number of constraints in terms of their and their rivals' 'capabilities' that derive from previous innovative activities; when competition is considered over a longer-time horizon, however, these ‘capabilities’ become endogenous to the competitive process as well.

Despite the absence of general results, game-theoretical models are certainly very useful to understand what fundamental strategic forces and incentives drive firms’ investments in R&D and what factors may lead to different types of dynamic competition (i.e., persistence of monopoly or action-reaction). When considering networks as dynamic systems game-theoretic models lend a rich flavor to the behavioral aspects of such systems.

Endogenous growth models, pioneered by Romer (1990), Lucas(1988), Aghion and Howitt (1992), Grossman and Helpman (1993) have tried to explain the economic forces driving the technological progress underlying economic growth. In these models economic growth, in general, derives from the growth of the stock of knowledge available to society, which in turn is the outcome of economic choices of profit-seeking agents. Earlier models developed in the 1960s were instead mainly focused on capital accumulation and did not treat technological change as an economic phenomenon amenable to economic analysis. Through these models we now know that growth does not generate high productivity or technological change, but vice versa, technological change generates high productivity and growth.

Although these models are based on the equilibrium methodology typical of the neo-classical approach they are inherently dynamic. The evolution of the stock of knowledge is explained as either being the product of unintentional learning-by-doing activities or the outcome of R&D investments undertaken by profit-maximizing agents and rewarded by the monopolistic rents that the innovator is able to harvest. Network externalities involved in the production of new knowledge usually play an important role in these models, and in particular knowledge spillovers by which the knowledge generated by an agent enhances the technological capabilities of other agents or their ability to discover new knowledge.

Having said that we first consider game-theoretical microeconomic models of innovation of the industrial economics provenience.

Industrial Economics

Within the new ‘industrial organization’ (Tirole, 1988), the game-theoretical literature on innovation which was developed vigorously in the 1980s and 1990s addresses the strategic interaction that characterizes R&D investment decisions in oligopolistic industries. These models provide the fundamental building blocks of more recent endogenous growth models which cannot be properly assessed without reference to this earlier strand of work.

More recently, game-theoretical models have focused on the analysis of competition and innovation in markets characterized by network effects, on the analysis of cooperation in research and development activities and on the study of the relationship between intensity of competition and innovation.

GTMs on R&D

Game-theoretical models of innovation explain the strategic interaction that underlies firms’ investments in research and development activities. The focus on strategic behavior distinguishes itself from earlier decision theoretic models, e.g., Kamien and Schwartz (1980).

Game-theoretical models of innovation can be classified as belonging to one of three broad paradigms:

1. General Rivalry and Racing Models (GRMs)

These models capture the strategic interaction that can arise in those cases where firms pursue different paths of technological advance so that a discovery made along one path can be used irrespectively on whether a competitor innovates along another technological path. This class of models has originated by Dasgupta and Stiglitz (1980). There are strategy based classes of GRMs that are embedded in EGMs such as Etro (2004) or Segerstrom (2006) in which 'leader-follower' interactions drive the growth process.

These models of innovation investigate R&D investments in a setting where there can be multiple discoverers and firms do not compete to be the first to innovate. In these models symmetric competitors cannot prevent each other from getting equivalent improvements from spending equivalent amounts of resources in R&D.

For this reason, such models may explain R&D investments in those environments where firms pursue different research paths so that an advance made on a single path can be appropriated irrespective of rivals' advances on other research trajectories.

A typical general racing model model is structured in three stages: an entry stage, an R&D stage and a market competition stage. Firms first decide whether or not to enter an industry; active firms would then invest resources in R&D activities to decrease their marginal cost of production or increase the quality of their products and would finally compete in the marketplace.

R&D activities are described in terms of a deterministic production function that specifies the size of innovation according to the amount of cost reduction or quality improvement per unit of R&D investment. Each firm can reduce its costs or improve the qualities of its products independently of parallel innovation by rival firms, although it is usually assumed that a firm can indirectly benefit from a rival's R&D activity for the existence of knowledge spillovers.

Knowledge spillovers capture the notion that innovative knowledge produced by R&D activities may, to some extent, be non-excludable so that firms can, at least partially, appropriate the results of R&D activities of a rival. Cohen and Levinthal (1989) suggest that such learning may not be costless and that firms may have to invest in own R&D programs in order to access external knowledge and benefit from spillovers from innovating rivals.

GRMs have especially been employed to assess the extent of innovation achieved by the operation of market forces and how it compares to the socially optimal one, to discuss technological policy, and to investigate the relationship between market structure (i.e., concentration) and innovation.

However, GRMs assume that there are a large number of research paths and the relationship between R&D investments and rewards from innovation is essentially continuous. For this reason they are not well suited to describe market environments where innovation takes the form of a race and competition is for the market rather than in the market. Although this may be an appropriate description for technological change in some markets, in many others innovation is a more disruptive process that leads to winners and losers in very different competitive positions. In these circumstances, the rewards to R&D investments are discontinuous, performance is rewarded on the basis of the rank within the set of realized performances, and competition takes the form of rival firms racing to be first to innovate (Gottinger, 2006, Chapters 1-3).

2. Auction Models (AMs)

Auction models of innovation describe R&D competition in terms of rival firms bidding for a given process or product innovation. Early contributions include Vickers (1986) and Katz and Shapiro (1987).

AMs of innovation describe R&D competition in terms of rival firms bidding for a given process or product innovation. The innovation is allocated to the firm which makes the highest bid and the issue is to investigate the identity of the firm which would have the larger willingness to pay for the innovation. Patent races could serve particularly well on auction models (Klemperer, 2003, Chapter 2) as well as auctions for purposes of mechanism design (Mas-Colell et al., 1995, Chapter 23).

In general, the maximum bid that a firm would make is equal to the difference between the profits it would get if it were successful in the auction and those it would get were it not. For this reason, auction models of innovation emphasize the competitive threat as the incentive underlying competition in innovation.

Auction models have been widely used to investigate dynamic competition and whether market dynamics are characterized by persistent dominance by a technological leader or by Schumpeter’s ‘creative destruction’ whereby the incumbent is overtaken by some rival whose incumbency is itself of short duration.

In a single auction model, the issue of dynamic competition is investigated by considering a set of firms with different technological levels, i.e., marginal costs in the case of process innovation, and studying their incentives to purchase an innovation. Market persistence is the result of such competition if the winner of the auction is the technological leader, otherwise dynamic competition is characterized by action-reaction. Assessing market dynamics in a single-innovation auction model, however, may not help to understand dynamic competition because in these models firms do not take into account how the outcome of the current bid may affect the outcome of future auctions. These inter-temporal links, however, may be an important element to explain firms’ incentives to invest in R&D in some contexts where a sequence of innovative opportunities is in prospect. For this reason, the question of dynamic competition may be better posed in the context of sequences of innovation rather than of a single innovation. Auction models are intrinsically deterministic and this leads to the somewhat disappointing result that losers do not commit resources in R&D and that the process of technological advance is not really explained.

3. Stochastic Racing Models (SRMs)

These models capture the strategic interaction that arises when firms compete to be the first to introduce an innovation and the issue of timing is crucial to firms’ strategic interaction. Papers in this category include Harris and Vickers (1987), Tirole (1988, Chapter 10), see also Gottinger (2006, Chapters 3-6).

SRMs of innovation capture the particular interaction that arises in those markets where dynamic competition takes the form of a race to be the first to make a discovery. As opposed to deterministic auction models, SRMs explicitly consider the case where there is technological uncertainty. This can be captured, for instance, by considering a stochastic relationship between R&D effort and success.

As such, these models emphasize timing in the context of R&D competition and are well suited to describe those environments where competition is for the market rather than in the market. This nature of market dynamics can arise because technological change proceeds along a small number of research trajectories and there is competition to be the first to innovate because success by a firm would somehow prevent benefiting by subsequent innovators. The advantage of being the first to innovate is usually explained on the basis of the ability of the winner to harvest the value of the innovation by means of patent protection. More generally, however, harvesting may depend on factors other than patent protection, such as lead-time, access to complementary resources and secrecy.

In auction models it is the competitive threat that characterizes firms' incentives to invest in R&D activities. When randomness is introduced, however, both the profit incentive and the competitive threat are relevant in firms' strategic interaction.

In general, these forces need not be symmetric and have to be considered for each firm involved: one firm may face the greater competitive threat and another may have the greater profit incentive. Clearly, if a firm faces both the greater profit incentive and competitive threat it would undertake more R&D and hence be more likely to win the race. However, where there are asymmetric incentives, the nature of the outcome is not as straightforward.

The literature usually focuses on the case where the competitive threat exceeds the profit incentive for all firms involved in the race. However, it could be the case that the profit incentive for both firms exceeds the competitive threat and that competition in R&D takes the form of a waiting game (Baumol, 1993, Chapter 7.5). Here some firms may tend to behave as free-riders and invest little resources in R&D. This type of situation has been considered, for instance, by Katz and Shapiro (1987) who explicitly account for the possibility of imitation and licensing of innovation.

As with auction models, SRMs have also been developed in sequential form in which firms complete a number of R&D stages before competing in the market. Examples of classical models that fall in this category are Fudenberg et al. (1983) and Harris and Vickers (1987). In these models innovation is obtained only as the result of the completion of a number of stages.

A crucial issue that arises in such a setting is the nature of the innovative process and in particular whether each firm has to discover each technological step itself before it can move on to discover the next or whether instead the follower firm can compete directly with the leader for the new state-of-the-art technology.

The latter case may arise if knowledge spillovers are relevant so that a laggard firm has access to (but may not have right to use) the same technological knowledge of the technological leader. On the other hand, knowledge may be excludable to some extent or in-house R&D is very important to assimilate external knowledge so that the technological frontier has to be achieved independently by each firm.

The Grossman and Shapiro(1987), Harris and Vickers(1987) step-by-step models assume that each firm has to discover each technological step itself before it can move on to discover the next(see also Grossman and Helpman, 1993, Chapter 4).

We find examples of leapfrogging models where a follower is allowed at any time to compete directly with the leader for the new best-practice technology. In a step-by-step model firms can compete neck-to-neck or have different technological levels. In a leapfrog model firms are always asymmetric in terms of their technological ability. This distinction turns out to be crucial in the analysis of the relationship between competition and innovation, as recent contributions in the industry founded innovation literature below have shown (Gottinger, 2006, Chapters 4-5).

Each of these approaches describes R&D investments in a particular setting, which may be an appropriate depiction of reality in some circumstances and a less appropriate theoretical reference in relation to some other industries or to particular stages of a technology life cycle.

Despite the differences, however, game-theoretical models of innovation illustrate that there are two central economic incentives at the basis of firms' investments in R&D activities, whose relative importance varies according to the specific environment considered: the profit incentive and the competitive threat (or replacement effect).

The profit incentive, or stand-alone incentive, is related to the notion that allocating resources to innovative research and development would, if successful, increase a firm's profits. This incentive can be usually thought of as the incentive to invest in R&D of a firm that takes the innovation decision in isolation. In fact, it is equal to the difference between the profits that the firm would get if it innovates and those that it would get if it does not innovate, all else equal.

The competitive threat, or pre-emption incentive, arises when strategic interaction in innovation activities is considered: a firm not only has to take into account the benefits connected to innovation but also the possible loss of competitiveness where and when it does not innovate and a competitor does. This incentive reflects the threat posed by the existence of an active rival and can be thought of as the difference between the profits if the firm innovates and the profits if it is left to a rival firm to innovate.

In general settings, both the profit incentive and the competitive threat may play a role in shaping the nature of dynamic competition. This would depend crucially on the specific characteristics of the market environment considered and, in the literature, on the specific assumptions adopted in each model.

Market Structure and Innovation

A long-standing question in the economics of technological change has been the nature of the relationship between market structure and innovation.

The relationship between concentration of an industry and its rate of technological innovation is complex. Market structure may have an impact on the rate of innovation, but innovation is also an important factor that shapes market structure. In fact, it is necessary to recognize that the relationship between concentration and innovation is not a causal one: both are the endogenous outcomes of the operation of market forces and exogenous factors such as the nature of demand, technological opportunity, the conditions governing use, and some chance.

The classical point of departure in the economic analysis of the relationship between static market structure (i.e., concentration) and innovation is the work by Arrow (1962). Arrow considers the case where a cost-reducing innovation is exogenously available and investigates firms' willingness to pay for the innovation under different market structures. For a drastic innovation, Arrow's analysis shows that a firm that is already a monopolist would have lower incentives to innovate than a firm that is currently in a perfectly competitive environment, essentially because it would have the lower profit incentive.

On the other hand, innovation is certainly an important factor that affects market structure: innovation is a means by which a firm tries to escape the constraints imposed by competition. Studies in the Schumpeterian tradition have emphasized the importance of ex-post market power for firms' incentives to innovate. Some degree of market power is necessary for a firm to cover its R&D outlays: dynamic and static efficiency are somehow conflicting. This is a theme that has been well developed in the recent literature on endogenous growth.

Dasgupta and Stiglitz (1980) discuss the relationship between concentration and innovation in a general racing model. If there are exogenous entry barriers, an increase in the number of firms causes each firm to spend less on R&D in equilibrium, however total R&D expenditure increases with the number of firms. When entry is considered to be endogenous, one would observe more innovation in those industries that are characterized by a higher degree of monopoly power, although no causality should be imputed to this relationship.

The models discussed above consider the relationship between product market structure (i.e., concentration) and innovation. In the context of models where firms race to innovate another interesting question concerns the relationship between the number of firms that are part of the race and the pace of technological advance.

These are racing models of innovation where R&D expenditures are committed upfront (that is, the probability of success depends on the scale of the R&D activity) and shows that increasing the number of firms reduces the expected date of invention.

Reinganum (1982) considers a setting where firms can vary the research intensity but does not assume that the rate of expenditure is constant over time. Instead, firms may adjust R&D intensity in response to elapsed time and rival progress. Reinganum shows that, in this setting, when imitation is not possible, an increase in the number of firms increases the equilibrium rate of investment for each firm and decreases the expected time of innovation. When there is no full patent protection, the relationship is ambiguous and depends on the relative payoffs to the innovator and the imitators.

Dominance

In many industries characterized by long-term market dynamics competition may take the form of competition for the market rather than competition in the market. In these markets the issue is not whether more or less concentration is associated with faster technological progress but whether market dynamics would be characterized by persistent dominance of the incumbent leader or by action-reaction whereby incumbents are overtaken by some rivals whose incumbency is itself of short duration. The dynamic evolution of market structure depends on both abilities and incentives of the incumbent and the rivals to innovate. Game theoretical models are well suited to analyze the incentives underlying R&D investments and the resulting evolution of market structure. If we focus on economic incentives, and set aside differences in R&D abilities, persistence of monopoly or action-reaction can be related to the different profit incentive and competitive threat faced by the leader and the follower(s).

The profit of a successful incumbent who innovates is that of a monopolist firm, whereas if it were the entrant to innovate each firm would get the profit of a duopolist. Hence, the competitive threat of the incumbent can be measured as the difference between the profit of a monopolist firm and the profit of a duopolist firm. The incumbent's competitive threat, instead, is simply equal to the profit of a duopolist. This implies that the incumbent would have more incentives to innovate (i.e., the larger competitive threat) if, as it is normally the case, the profit of a monopolist is greater than the combined profits of two duopolists.

Katz and Shapiro (1987) show that for minor innovations, the industry leader will typically be the innovator, whether or not imitation and licensing are feasible. For markets where patent protection is strong, they predict that the major innovations will be made by industry leaders. But if imitation is easy, industry followers or entrants will make the major discoveries. Racing models show that when the first innovator captures a sufficiently high share of the post innovation market, then the incumbent firm invests less on a given project than does the potential entrant. This is because the incumbent has less incentive than the challenger to shorten the period of its incumbency.

Market dynamics have also been investigated in models that consider a sequence of innovations. Reinganum (1985) considers a sequence of drastic innovations and shows that market dynamics are characterized by a process that resembles Schumpeter's process of creative destruction: the incumbent invests less than each challenger in each stage. Vickers (1986) takes a sequence of non-drastic process innovation in the context of the auction model. He compares market dynamics under Bertrand and Cournot competition and finds that when the product market is very competitive then there is increasing dominance, but when it is not very competitive there is action-reaction. When innovation is drastic, then market dynamics are characterized by increasing dominance. Reinganum (1989) observes that the differences in the results obtained in different models can be ascribed to the different roles that the profit incentive and the competitive threat play in racing and auction models. In a deterministic model the incentive to pre-empt (larger for the incumbent) dominates the firms' decision. When success is stochastic, however, the threat from the rival innovating is less acute. In the case of drastic innovations, the competitive threat is the same for both firms and it is the profit incentive (which is larger for challengers), next to diverse non-monetary incentives, that determines the level of R&D investments. The relevance of the profit incentive extends to the case for some non-drastic innovation.

Competition Intensity and Innovation

Market structure is often associated with the concept of competitiveness: usually, a high level of concentration in an industry is interpreted as weak competition. This view is based on a symmetric Cournot model, where price-cost margins are higher as the number of firms increases. However, it is preferable to disentangle the notions of market structure and toughness of price competition as, among others, in Sutton (1998). For Cournot vs. Bertrand in theories of oligopoly behavior we can derive the following results (Shapiro, 1989)

In a simple world with homogeneous firms, toughness of price competition can be considered as being related to the level of price-cost margins given any level of concentration. Hence, a differentiated Bertrand market would be considered to be more competitive than a differentiated Cournot market because price-cost margins would be lower in the former, for any level of concentration. As a result, more competitive markets may allow fewer firms to profitably survive.

When intensity of competition is low it is the follower that would be the next innovator, whereas when intensity of competition is large, it is the current technological leader that is likely to innovate. When it is the follower that innovates, tougher competition implies lower profits and hence lower incentives to innovate. However, when it is the current leader to innovate, an increase of toughness of price competition would further increase the profits related to his technological leadership and hence would increase the value of innovation for the firm. Hence, the relationship between toughness of price competition and innovation would be inversely-shaped.

Other studies that investigate the relationship between toughness of competition and innovation fall within the endogenous growth literature (Aghion and Griffith, 2005).

R&D Cooperation

Increasingly, research joint ventures (RJVs) have become a widespread form of industrial cooperation (Gottinger, 2006, Chapter 9). In parallel, the study of the effects of R&D cooperation has also emerged as an important research topic when considering industry policy. Cooperative R&D ventures can take many forms, ranging from R&D joint ventures, to cross- licensing agreements and various informal types of technology trading or information-sharing agreements.

RJVs themselves can cover a variety of arrangements. One type of RJV is the traditional joint venture, in which two or more parties create a separate entity in which they all have equity interests to conduct well-defined R&D projects for their benefits. Another type is the research consortium. A third form of RJV is the venture capital investment by firms in a stand- alone start-up company.

Economists have investigated the extent to which RJVs might allow firms to internalize spillovers, coordinate their research activities and achieve higher R&D efficiency. Cooperative R&D is thought to be socially beneficial for several reasons:

RJVs can alleviate the under-provision of R&D effort that results from technological spillovers and other sub-optimal R&D decisions.

Cooperation can lead to greater dissemination of R&D results. It can improve R&D efficiency through good research design and information sharing, avoiding needless duplication of resources. Cooperative R&D enables the firms to share risks, to exploit synergies, pool different complementary assets, and exploit increasing returns to scale in R&D. It can enable firms to overcome a cost-of-development barrier impenetrable to any one of them alone.

Adding to these effects, an important issue that is not very developed in the industrial organization literature but is emphasized in the evolutionary economics and the management literature on inter-firm cooperation is that cooperative ventures may be an important means by which firms exchange specialized know-how and tacit knowledge. Countering these social benefits is the fear that firms participating in cooperative R&D might use the venture to engage in anti-competitive practices or that they might free ride on each other.

5.3. ENDOGENOUS GROWTH MODELS

Technological change is a fundamental driver of economic growth. However, only recently have the economic forces that underlie technological progress in a dynamic economy been investigated. Advances in this field have been associated with the development of endogenous growth theories, one of the most fertile grounds for economic research in the last 30 years. The relevant research agenda has been set on the study of economic growth as the result of knowledge production — we shall focus on technological knowledge in this review — which is explained as the outcome of economic decisions.

Before the development of endogenous growth theories economic models of growth were essentially models of capital accumulation developed on the basis of Solow's (1956) seminal contribution. These models suggested that the simple accumulation of physical capital cannot sustain long-run economic growth as long as its marginal productivity is decreasing with diminishing returns to scale. Long run economic growth had to be attributed to the effect of some other factor, exogenous to the economic choices analyzed in these models, such as technological change.

In fact, even in older exogenous growth theories that studied capital accumulation, technological change was acknowledged as one of the major factors that explain the ability of an economy to grow in the long run. Certainly, advances in technology are a desirable element in any explanation of growth:

> "…a story of growth that neglects technological progress is both ahistorical and implausible" (Grossman and Helpman, 1993, 26).

Economic history studies have also examined the relationship between knowledge and economic growth. Nevertheless, for a long time the forces underlying advances were not explained on the basis of economic incentives but were essentially related to the exogenous evolution of scientific knowledge and its technological applications.

Endogenous growth theories or new growth theory have certainly contributed to fill this gap. These theories are based on product differentiation and imperfect competition, economies of scale and increasing returns instead of perfect competition and diminishing returns put forward by neoclassical theory. New growth theory not only emphasizes knowledge accumulation and diffusion, in the form mainly of human capital (Lucas, 1988) or technological innovation (Romer, 1990; Aghion and Howitt 1992), as fundamental drivers of economic growth but also the economic decisions that underlie both.

As such, this review focuses on the second generation of endogenous growth models that consider technological innovation as well as human capital accumulation. In general, these models relate economic growth to innovations that improve firms' productive efficiency and/or lead to the production of new or better intermediate and consumer goods.

Innovations result from new knowledge produced as the outcome of firms' learning-by-doing and/or intentional R&D activities. In general, R&D models of endogenous growth have drawn heavily from the earlier game theoretical literature on innovation and extended those analyses to more dynamic settings. As such, they have developed the analysis of long-standing questions investigating, for instance, the relationship between the intensity of competition and innovation and the nature of market dynamics. More recent economic analysis in this field has offered interesting new insights on the relationship between competition and innovation such as those in Aghion and Griffith (2005) which make only sense through network analysis.

Knowledge generated by learning-by-doing can be a significant engine of economic growth and models of endogenous growth with technological progress driven by learning-by-doing have been studied notably by Aghion and Howitt (2008).

Learning-by-Doing

Learning-by-doing, since the seminal work of Arrow (1962), Solow (1997) has been considered an important source of technological knowledge. Firms can produce new knowledge without investing in institutionalized research and development activities but rather, as a by-product of their usual production activity. Indeed, experience has been shown to be an important source of technological improvements in many industries such as artificial fibers, semiconductors and memory chips.

Generally, in these models learning-by-doing is the unintentional byproduct of production or investment activities and it is usually considered to result in technological knowledge, equivalent to incremental and continuous technological change. The existence of knowledge spillovers is in fact generally acknowledged: a firm can, in part at least, benefit of the knowledge generated by the learning-by-doing of other firms.

We could also observe settings where new technologies are discovered through R&D activities but are initially inferior to the older technologies they seek to replace. Experience, however, generates incremental improvements over time, which allow the new technology to supplant older technologies. These incremental improvements are only achievable to some extent, in the sense that there is a bound to the technological advance that can be generated by learning-by-doing. This hybrid model emphasizes that inventive activity and production experience are complementary forces in driving technological change.

A similar enlarged network externality effect of ‘learning-by-doing’ for growing international trade could result in ‘learning-by-exporting’ with positive and cumulative benefits for economic growth (van den Berg, 2001, Chapter 9.3).

R&D Models

Endogenous growth models have also investigated economic growth as driven by the accumulation and diffusion of knowledge generated by Research and Development activities, undertaken by profit-seeking firms.

These models investigate the process of discovery of new or better products. Accordingly, a distinction can be drawn between those models that treat innovation in terms of the development of new varieties of products (horizontal expansion) and those that consider sequential improvements of the quality of existing products (technology deepening). They are intrinsically related to network economies that link innovation effort to measurable quality improvement resulting in economic welfare enhancement (Gottinger, 2003, Chapter 7).

The major difference between these two types of models is that the latter capture an important characteristic of the innovative process, namely that new inventions entail an element of destruction because they make old technologies or products obsolete. For this reason, early endogenous growth models with sequential improvements in the quality of products are called Schumpeterian.

Despite these differences however, the two types of models share the same fundamental structure. In both models, firms invest resources to acquire the exclusive right to manufacture a new product and it is usually assumed that R&D activities generate knowledge spillovers that benefit other research firms. The production of new knowledge and its diffusion drive economic growth in the long run.

Product Variety

Romer (1990) provides the seminal model where technological progress is captured in terms of an increasing variety of products. The increase of the variety of intermediate or final goods is reflected in higher productivity or utility and hence economic growth. New varieties of products are discovered as outcomes of R&D investments and remunerated by the rents that accrue to the successful innovator who can benefit from patent protection (or, more generally, of other factors such as lead time, secrecy, access to complementary resources).

In these models, it is generally assumed that R&D activities lead both to new knowledge that can be appropriated by the innovator and new general knowledge that cannot be appropriated but instead contributes to a stock of publicly available knowledge that facilitates further discoveries.

Not only is the production of new knowledge necessary for economic growth, but also its diffusion.

A distinguishing feature of Romer's (1990) growth model is that it studies horizontal product innovations that involve no obsolescence: new goods are never close substitutes for existing goods (i.e., each new product finds its own horizontal niche).

There are some industries where innovation is fundamentally directed at the introduction of new varieties of products rather than at improving products' qualities.

Product Quality

Aghion and Howitt (1992) examine growth as driven by industrial innovations that improve the quality of products.

Vertical innovation models of this kind introduce, relative to models with expanding varieties of products, the consideration of the process of creative destruction: a new innovation makes, to some extent, old products obsolete and negatively affects the economic rents of previous innovators. For this reason, models with qualitative improvements are also called Schumpeterian because they embody Schumpeter's idea of creative destruction by emphasizing the process by which new products (and innovators) displace old products (and innovators).

Endogenous growth models with vertical innovations have very elaborated micro-foundations describing the nature of market dynamics, which are similar to and develop further earlier game- theoretical models of innovation. As such these models have developed the study of dynamic competition.

Aghion and Howitt (1992) study a model of economic growth based on creative destruction. Innovation consists of a higher quality intermediate good that can be used to produce output more efficiently than before. Research firms invest resources in R&D activities that stochastically determine the time of the new innovation; the arrival rate of new innovations depends only upon the current flow of input to research. Firms' R&D investments are motivated by the prospect of monopoly rents that can be captured when a successful innovation is patented. However, unlike the horizontal innovations considered in Romer (1990), these rents would last only until the next innovation occurs, when the current innovation would become obsolete.

Aghion and Howitt consider the cases of both drastic and non-drastic innovation. In the first case the new innovator is unconstrained by potential competition from the previous patent; in the latter case, such a constraint is binding.

In the case of drastic innovations market dynamics take the form of a continuous process of action and reaction whereby at each point in time the market is dominated by a monopolist whose incumbency is short lived. In theory, a new innovation could be introduced by either the current monopolist or an outside research firm. The value of innovation is the expected incremental present value of the flow of monopoly profits generated by the innovation over an interval whose length depends on the occurrence of the next innovation. Because this incremental profit is lower for an incumbent monopolist, only outside research firms would innovate and the monopolist chooses to do no research.

The reason for this result lies in the fact that, given the assumption and the structure of the model, firms' incentives are driven only by the profit incentive, which in the specific case is always larger for an external firm. This is the replacement effect emphasized by Arrow (1962) whereby a monopolist would have lower incremental profits from innovation than a perfectly competitive firm. In the case of drastic innovation the competitive threat does not play a role because the flow of profit is independent of the identity of the innovator. Innovations are non-drastic if the previous incumbent could make a positive profit when the current one is charging the unconstrained optimal monopolistic price. In the case of non-drastic innovation the competitive threat may potentially play a role in shaping equilibrium investments in R&D activities. However, Aghion and Howitt (1992) assume that the monopolist chooses to do no research also in the case of non-drastic innovations, i.e., assume that the efficiency effect is smaller than the replacement effect. As a result market dynamics takes the form of continuous action/reaction.

Finally, Aghion and Howitt (1992) consider the case in which firms can affect not only the frequency but also the size of innovations. Their finding is that innovations would be too small if they were drastic; in the non-drastic case, the tendency to make innovations too small is in part mitigated by the incentive for innovators to move away from their competitive fringe. Grossman and Helpman (1993, Chapter 4) have built on the model by Aghion and Howitt(1992) to consider an economy with a continuum of, and not only one, intermediate goods. Each product has its own quality ladder and entrepreneurs target individual products and race to bring about the next generation. In each industry, success occurs with a probability per unit of time that is proportional to the total R&D resources invested to improving that product. The authors allow free entry in the race for the next generation of products and assume that potential entrants can learn enough about the state of knowledge to compete to produce the new state-of-the-art quality. These assumptions imply that, without a cost advantage, industry leaders do not invest resources to improve the quality of their own state-of-the-art products. This is because of the replacement effect whereby the leader would obtain an incremental flow of profits that is less than that of an external firm.

Aghion and Howitt (1998, 2008) introduce an element of heterogeneity in innovative activity that captures the distinction between fundamental and secondary research. They first investigate this heterogeneity by considering fundamental research as deriving from R&D activities and secondary innovation from learning-by-doing. Each innovation resulting from research consists of a potential new product, and learning-by-doing leads to improvements of the quality of goods that have been invented. They also consider the possibility that learning-by-doing contributed to a general stock of knowledge that creates new opportunities for research as well. Moreover, Aghion and Howitt (1998) consider both the case where learning is not taken by the firm but is shared by all firms and where the quality enhancement of learning is fully internalized. Also they consider the case where the distinction between fundamental and secondary innovations is captured in terms of research versus development. A feature of Aghion and Howitt's (1992) influential model, shared by most early models of endogenous growth with quality ladders and leapfrogging, is that more competition (i.e., higher elasticity of demand for each firm) is associated with less innovation and growth, a very Schumpeterian result. The reason for this conclusion is that these models capture only an increase of ex post product market competition and not of ex ante competition as well. Some recent studies modify this feature of earlier models and reach a different conclusion on the relationship between competition and innovation: tougher competition may increase the incentives of firms to innovate in order to escape from a more competitive market state.

The substantial difference of these models compared to earlier Schumpeterian models lies in the fact that, by considering the possibility that the incumbent innovates, it is the incremental profits of each firm that drive R&D investments and the impact of product market competition has to be evaluated both ex ante and ex post. Tougher product market competition might reduce a firm's pre-innovation rents by more than it reduces post-innovation rents, hence it may increase R&D firms' investments.

In these models, technological progress does not involve leapfrogging. Hence, a typical industry can be in either of two states: it can be a leveled industry where firms compete neck-to-neck or an unleveled industry where one firm is technologically ahead of the follower.

The effect of an increase in product market competition would have to be evaluated considering the impact in both leveled and unleveled industries. The impact is actually twofold: on the one hand the degree of product market competition affects equilibrium R&D investments in each of the possible state of the industry (leveled and unleveled); additionally, product market competition has an impact on how often the industry will be in one of either states. In a leveled industry increased competition would spur innovation because it would increase the incentives of firms' to get ahead of the rival (the new 'escape competition effect'); in an unleveled industry, increased competition would reduce the incentives to innovate for the classic Schumpeterian effect. This reasoning is in Aghion and Griffith (2005), which leads to the conclusion that the relationship between intensity of competition and innovation has an inverse shape. Consider the case of very large innovations, in which case the leader never innovates and the largest gap between the leader and the laggard is one technological step. When there is not much product market competition, there is hardly any incentive for firms to innovate when they compete neck-to-neck, and the overall innovation rate will be highest when the industry is in an asymmetric state. Thus the industry will spend relatively more time in the leveled state where the escape-competition effect dominates. An increase in product market competition would then result in larger incentives for the firms to innovate. Hence, if the degree of competition is initially low, an increase will result in a faster innovation rate. When initial product market competition is high, there is relatively little incentive for the laggard firm in an unleveled state to innovate and a relatively large incentive for a neck-to-neck firm to leave the leveled state. The consequence is that the industry will spend most of the time in the unleveled state where it is the Schumpeterian effect that dominates. Tougher product market competition reduces the Schumpeterian effect so that when the degree of competition is initially high, an increase may result in a slower average innovation rate.

5.4. Innovation in Evolutionary Economics

Evolutionary models draw their theoretical foundations on different sources and particularly on biological models, evolutionary game theory and institutionalism in order to capture essentially dynamic and irreversible processes.

In view of the description of technical change the models embrace more broadly and systematically innovation and knowledge (Foray, 2006). The modelling approach avoids the pursuit of full rationality and equilibrium, and also incorporates a consideration of institutional factors. Though these themes trace back to earlier work — in particular, the Schumpeterian perspective, which places emphasis on dynamics and competition through innovation — dynamic theorizing in evolutionary economics is often seen to start with the theoretical construct of Nelson and Winter (1982).

The literature tends to place emphasis on innovation and selection from variety as the force of economic development. Of particular importance is the need to see innovation as more than simple technological development (and much more than the production of knowledge as a public good) and to analyze economic development through technological and organizational improvements (which lead to the introduction of new products, inputs and techniques).

In evolutionary theories innovation is in fact defined more broadly in a societal and social context than in business related models and as a multidimensional phenomenon. Innovation is not only a reduction in production costs or the commercialization of new products but also a more efficient reorganization of a firm's activities, the adoption of more advanced procedural and 'management science' techniques, access to new markets, etc.

This richer definition emphasizes that innovation extends outside the realm of technology in two ways: many innovations are not technological but involve other aspects of a firm's activities such as its organizational structure; moreover, even technological innovations often affect non-technological activities of a firm: technological and organizational changes are often importantly intertwined.

Another important difference with traditional models of innovation is that evolutionary models emphasize the tacit and specific nature of knowledge, as opposed to the idea of knowledge as information that somehow underlies common models of innovation. In terms of the process that leads to the discovery of new products, evolutionary models describe the importance of learning, and of organizational learning often embodied in organizational routines as a source of knowledge.

Evolutionary economics can be looked at both from a microeconomic and macroeconomic perspective. The microeconomic aspects of evolutionary economics are mainly concerned with how innovation arises at the firm and market level. In contrast, the macroeconomic strand of evolutionary economics places innovation into the social and economic context and examines the dynamic process through which the innovation systems generate innovations through a complex and reciprocal relationship network connecting different components of the system.

Innovation and Competition

A key feature of evolutionary economics is its emphasis on competition as a process rather than an outcome. Adopting a more dynamic approach, evolutionary economics is in part associated with the use of analogies to evolutionary biology to explain economic growth and the process of competition. At a very basic level, the notion of competition through natural selection associated with Darwin seems similar to the process by which economic competition selects more fit (efficient and profitable) firms at the expense of less fit firms. In this context, Mokyr (1990) defines the Darwinian model "as a system of self-reproducing units (techniques) that changes over time."

Excessive variety in a Darwinian system implies that the actual number of techniques exceeds the number sustainable in the system. This results in a selection process.

Mokyr (1990) argues that selection can operate through three channels. The first one is the standard neoclassical mechanism where techniques are chosen in the event they maximize an objective function incorporating supply and demand side consideration as well as externalities. Second, there is inertia in the system that eliminates useless techniques. Finally, and more importantly, selection not only happens in the market but also at the social level.

Inherent in this model of competition is the association between competition and experimentation.

Thus economic development and innovation can be seen as a combined effect of selection (via competition) from a variety of competing routines and practices as well as the more endogenous process of agents seeking improved routines and practices. The role of 'dynamic' competition in this regard is quite different from its 'static' force in constraining prices.

This emphasis on competition as a selection process is supported by a quite separate strand of empirical literature that seeks to decompose industry-level productivity growth to examine the 'microstructure' of this growth. Such studies use panel data of firms (or plants within multi-plant firms) to examine the extent to which economic development at the industry level is driven by productivity improvements within firms ('internal restructuring') versus expansion (and entry) of high-productivity firms and contraction (or exit) of low-productivity firms ('external restructuring').

Note that such studies focus more on technological and organization improvements than drastic innovations (major new lines of products) that would not be fully captured in the productivity studies. But it is clear to see how evidence of internal versus external restructuring, in the context of industry-level economic development fits in with the evolutionary concepts of competition as a selection process.

Industry Structure and Innovation

Schumpeter is often associated with the hypothesis that large firms with market power are more innovative than small ones. Can we explain why larger firms may be more innovative than smaller firms. On the one hand, if firms use retained profits to finance R&D, large successful firms will have advantages over smaller firms. On the other hand, a large firm is able to derive greater benefit from finding a better ‘routine’ because it can put it to use across a larger number of units of production. Such analysis would suggest that large firms are at a competitive advantage, although Nelson and Winter are reported to see this effect counteracted by large firms (with more market power) having a higher profit target (price/cost ratio) and this may provide some restraint in these dominating smaller firms.

These models also suggest that if large firms are aggressive in pursuing their advantage over small firms, and there is in turn a tendency towards concentration, large ‘imitating’ firms may do well at the expense of smaller innovators. Thus where technological progress is endogenous, productivity may be hindered compared to a situation of a more varied industry structure.

An alternative way of considering whether large firms are more innovative than small firms is to focus more on heterogeneity in ability. To the extent that competition leads to the growth of innovative firms, we might simply expect causality to run the other way: it is not that large firms are inherently better innovators than smaller firms, just that the better innovators will be successful and hence grow in size. Theories based along these lines would therefore be treating innovation as exogenous to firm size, with causality running from the former to the latter.

This exposition between exogenous and endogenous innovation is valid more generally than firm size. Nelson and Winter (2002) draw a distinction between industries that are characterized by ‘science-based’ innovation and industries that are characterized by ‘cumulative’ innovation. Science-based industries are considered to be those where the thrust of innovation comes from R&D activities outside that industry. For example, firms in such an industry might benefit principally from external scientific developments and innovative activities by their suppliers. By contrast, in cumulative industries, innovation stems from R&D activity within each of the firms in that industry though in some industries this distinction is artificial: as in the early German chemical industry and in more recently developing biotech and semiconductor based industries where technology entrepreneurs came from academia for startups with experiences fed back to academia for viable research.

5.5. Systems of Innovation and the Role of Institutions

The emphasis on the notion of ‘systems’ of innovation suggests that innovation is not a phenomenon which results from the activities of isolated innovative firms, but the result of complex important interactions between firms and other organizations and institutions. Each innovative activity by a certain organization involves an element of reliance on external sources and innovation in the economy results from the complementary contributions of different organizations in a particular institutional environment.

Different definitions of the concept of innovation systems have been offered, and, despite the inevitable differences, all point towards the importance of the notion that invention, innovation and its diffusion result from interactions between different complementary organizations and institutions.

It can be thought of as a system which creates and distributes knowledge, utilizes this knowledge by introducing it into the economy in the form of innovations, diffuses it and transforms it into something valuable, for example, international competitiveness and economic growth.

In general a system of innovation is defined in terms of its components and the relations through which the components are linked with one another.

The elements of the system of innovation are organizations and institutions. Institutions exert a considerable influence on organizations, and simultaneously, organizations also impact on the institutional environment. Institutions may give rise to new organizations as well as being the origin of new institutions.

On the one hand it is argued that the element of nationality may derive from different factors such as the national focus of technological and other policies, laws and regulations that have an impact on the innovative environment. Because the education and research system, public infrastructure, laws and financial institutions keep some of their national characteristics, differences across different national innovation systems are likely to persist even in the longer run.

On the other hand, the boundaries of systems of innovation in different countries are getting blurred with the rise of international firms present in several countries and increasingly intensifying cross-country inter-firm connections to share knowledge and innovation.

It is widely recognized in the literature that there is a dilemma in national systems of innovation. This dilemma is whether to create incentives for innovation or whether to foster diffusion of innovation. At the heart of this problem is the patent system that encourages innovation by allowing for reaping the benefits and which at the same time hinders diffusion.

Mokyr (1990) also suggests that some institutional arrangements may hinder innovation. He points out that self-organizing systems become the cornerstone of evolutionary economics. He also emphasizes that resistance in society might have not only a stabilizing effect but can also act as a break on innovation. Indeed, high resistance can slow down or even eliminate the introduction of innovations or even deter innovation activity. Opposition against new innovations might come from both owners and employees.

In a series of seminal studies, North (1990, 2005) underscores the importance of social and economic institutions, which are likely to affect innovation and hence long-term economic development. According to North, institutions form the incentive structure of a society, and as a consequence, the political and economic institutions are the underlying determinants of economic performance.

The institutional environment provides a general framework for social and economic interactions.

Institutions fall into two categories. The first category can be referred to as formal constraints and is composed of rules, laws, constitutions, property rights, whereas the second one is labeled as informal constraints and is made up of customs, traditions, behavior, conventions, codes of conducts, etc.

In North (1990), the evolution of institutions is analyzed in a historical perspective. It is argued that there are some primitive institutional settings that are most unlikely to experience changes triggered by pressures coming from the inside. In three types of exchanges under consideration, namely tribal society, regional economy and long-distance caravan trade, learning, and hence the accumulation of knowledge and skill will not lead to changes, simply because innovation is perceived as something endangering the system's survival.

All in all, North emphasizes that productivity gains were not due to technological progress but rather the consequence of organizational innovations and changes in the institutional environment. In view of increasing returns paths of network economies, as discussed in Chap.4. North(1990, Chapter 11) calls for institutional change through 'institutional increasing returns' reinforcing the dynamics of innovation.

According to Mokyr (1990), two types of inventions can be distinguished. The first is 'macro invention' and describes inventions of paramount importance for the economy as a whole. The second is 'micro invention' and consists of small steps by which already extant techniques are ameliorated. Mokyr (1990) points out that macro inventions

> "… do not seem to obey obvious laws, do not necessarily respond to incentives, and defy most attempts to relate them to exogenous economic variables. Many of them resulted from strokes of genius, luck or serendipity."

Mokyr notes that Britain was by no means in a better position in the realm of macro inventions compared with France but the institutional and economic environment were more conducive as regards micro inventions.

5.6. Innovation and Competitive Advantage

In considering the strategies employed by innovative firms, it is useful to draw a distinction between two broad categories: how a firm gains a competitive advantage over rivals; and how a firm can protect and maintain that competitive advantage.

Value creation: how firms gain a competitive advantage.

The first strategy focuses on how firms can create more value, which can be achieved in principle through cost reduction as being part of efficiency improvements; increasing quality and value of existing products and services via differentiation relative to competitors; and developing new products and services.

At the center of value creation is the notion of competitive advantage introduced by Porter (1990, 23). Porter's view is this:

> "Competitive advantage is at the heart of the firm's performance in competitive markets. After several decades of vigorous expansion and prosperity, however, many firms lost sight of competitive advantage in their scramble for growth and pursuit of diversification. Today the importance of competitive advantage could hardly be greater. Firms throughout the world face slower growth as well as domestic and global competitors that are no longer acting as if the expanding pie were big enough for all."

In accordance with Porter (1990), competition is a continuous quest for competitive advantage. External and internal sources of competitive advantage can be identified. External sources such as changing consumer patterns or changes in technology are exogenous to firms. Since firms are heterogeneous in their resources and capabilities, they are not equally able to adapt to such changes, and thus some firms are able to realize competitive advantage through more timely and efficient responses to exogenous factors.

The internal source of competitive advantage is essentially the capability to innovate, by which we mean inventive effort, rather than adaptive or imitative behavior.

Cost Leadership

As highlighted above, value to customers can be created by means of decreasing costs, through which a firm is able to reduce prices relative to competitors.

At the extreme, focus on this strategy can be described as thriving for cost leadership, which refers to the situation when a firm has the capability to produce similar products at significantly lower cost than its competitors. There are several routes through which a firm can achieve cost advantage, for example:

Economies of scale and scope yield benefits to the extent that an increase in production volume will entail a reduction in unit cost.

Economies of learning describe experience-based learning. The more complex the process technology or the product, especially in terms of tacit knowledge (i.e., that which cannot be codified and transferred) the larger the scope for benefits from learning-by-doing. This provides means for established companies to experience cost advantage over new firms.

Enhancing process technology and process design allows for greater productive efficiency and lower production costs for the same output.

Similarly, enhancing the organizational efficiency and routines (for a given technology) will increase productive efficiency.

Reducing input costs allows for lower unit costs for the same level of productivity.

Because of the use of different suppliers, geographical differences in input prices and different bargaining power, different firms may pay a different price for the same set of inputs. Augmenting capacity utilization: In the presence of high fixed costs, unit costs are to be decreased by increasing capacity utilization. On the other hand, overcapacity resulting in overtime pay, premiums for night shifts and rising maintenance costs also increase unit costs.

Product Quality and Differentiation

Changes to product quality (in particular quality-price trade-offs) and product differentiation compared to rival products represent the second facet of value creation.

A firm can set out to differentiate its product according to the tangible and intangible aspects of the products it manufactures. The former concerns physical characteristics and performance of the product, while the latter relates to perceived, mainly social and psychological faculties.

The nature of competition in a market will be associated with whether firms strive to gain competitive advantage through innovation in physical attributes of the products they supply, or the intangible characteristics. For example, where an industry is mature and there appears little scope for changes in product innovation and cost leadership, competitive advantage may be achieved through a successful branding exercise that provides differentiation relative to rivals.

Product Innovation

While changes to the quality of a product represent some degree of product innovation, innovative behavior is likely to be greater with regard to the introduction of new products and services that have features to mark them as distinct from current products. A strategy of product innovation may have the effect of creating new markets (in competition policy terms, the introduction or products for which there are no good substitutes on the demand-side) but may also, though not necessarily, destroy old markets (as consumers cease consumption of previous generation of products in favor of the new products).

When it comes to assessing how to exploit an innovation to the maximum, that is to maximize profits related to the innovation, firms can choose from several options depending on how much risk can or are they willing to take, and thus how high a return do they expect; how many resources can or are they willing to put at disposal when exploiting the innovation.

Risks are high in emerging, innovative industries. First, firms have to face technological uncertainties (such as the direction technology takes in and which technical standards are established at the industry). Second, there are substantial market uncertainties. In other words, it is difficult to predict the potential size and expansion of the market.

The strategy of internal commercialization means that the firm commercializes the new product that is, the output of the innovation, without any external productive help.

This is the highest level of involvement that goes hand in hand with the heaviest investments and therefore the highest risk. Nevertheless, the firms can have the whole process under its control. In doing so, the resource requirements are substantial, which implies that the firm is expected to be fairly large to be able to bear the burdens.

The decision to start a joint venture still implies a great deal of involvement. By contrast, it also helps to share the cost and the risk of the investment. In addition to that, the participating firms can pool their resources together and thus benefit from synergies. On the other hand, however, they run the risk of potential disagreement and possible differences in corporate culture.

Strategic alliance represents a middle way between total and very small involvement. Risk and resources can be shared. Outsourcing is suited to both large and small companies so as to enable access to outside resources without too much commitment.

Licensing is typically but not exclusively employed by small firms, which do not possess the necessary resources to exploit the innovation on their own.

Competitive Advantage

The second strategy can be understood as efforts to protect and maintain a competitive advantage that has been achieved through value creation. In particular, it concerns efforts firms may take to prevent imitation by rivals, since in competitive environments, imitation is a crucial force that undermines the competitive advantage of a firm over time. Such behavior may be both legal (e.g., patenting) and illegal (e.g., certain types of anticompetitive behavior that inhibit rivals access to a market).

Once a firm has been able to gain competitive advantage in some respect, it will have incentives to protect and maintain this. In particular, where a firm has benefited relative to competitors from an important innovation it will be keen that rival firms do not undermine this advantage through imitation (Jaffe and Lerner, 2004).

One way in which firms achieve such protection is through various legal institutions put in place for this purpose (e.g., patent policy).

The Use of Legal Institutions against Imitation

Patent policy provides a firm with protection against direct imitation of certain types of inventive innovation, for a period of time, subject to disclosure of information relating to the innovation. Similarly, copyright protection prevents the direct copying of new content. In both cases the aim of the legal protection is to provide incentives for the innovation to take place, exactly because in the absence of this protection such incentives for such innovation would be mitigated by the threat of direct imitation and risk of immediate erosion of any competitive advantage accruing to the innovation.

Nonetheless, there are ways in which patent policy may be used by firms against the interests of consumers. Preemptive patenting describes the case where firms gather a series of patent around an initial innovation, not for the purposes of using those patents in the introduction of new products, but rather to stop other firms form inventing around the original innovation. Thus the new patented products or technologies are neither used by the patentee nor licensed to other firms, but their value derives from the fact that rivals cannot use them rather than the use the holder makes of them.

Alternatively, where a particular product or service protected by patent represents a bottleneck to a range of related markets, the patent holder may be able to use its protection in the supply of the patented product to enjoy competitive advantage in the supply of products in these related markets.

In some sense, there is overlap between creation of competitive advantage through product differentiation and maintenance of this advantage through branding. Where firms can link an innovative product to a successful brand image, and protect that brand from imitation (e.g., via trademark law), the advantage from the initial innovation can be maintained in the face of imitation as to the products physical characteristics. For example, in the pharmaceutical industry, successful creation of goodwill, reputation and brand recognition allow firms to enjoy some (albeit less) competitive advantage over a drug that they have created once the patent has expired.

Timing of Innovation and Life Cycles

Having set out firm strategies according to two broad categories, we discuss below issues of timing. Two questions arise: how is the success of innovation, and thus the optimal innovation strategy, dependant on timing. How do product 'life-cycles' arise more generally?

First-Mover advantage

An essential question in innovative markets is to choose a strategy whether to enter early in the market and invest heavily in research and development in the hope of future profits and market leadership or to adopt a wait-and-see policy and to follow the technology leader by copying the innovation. The benefits of a first-mover strategy relies on three conditions:

Length of the lead time (i.e., the time over which the innovation is protected from imitators and followers) either through IP rights or less formally.

The extent to which complementary resources are required while exploiting the innovation. If the need for complementary resources is relatively low, the initial investment and therefore the risk is lower. This may stimulate firms to try to be the first-mover.

The possibility to set a standard is also likely to give an impetus for first-movers. The reason why standards are set is closely connected with positive network externalities (discussed further below in the context of the economics of networks). The higher the number of users of the standardized product, the more valuable the product to individual user. Positive network externalities have three major sources:

(i) users of the product are connected with each other through networks (such as transportation and the Internet);
(ii) the provision of which complementary products are available (for example, software applications for Windows); and
(iii) the switching cost between different networks: the larger the network the customer uses, the lower the costs to switch to another network (for example, mobile phone networks).

Generally, technological standards tend to result in ever increasing positive network externalities. If a technological standard appears to be dominant in the industry, more and more users are likely to choose this one that may easily lead to winner-takes-all situation. This is one reason why innovative firms try hard to establish their technology as the industry standard.

Shapiro and Varian (1999) identify three possible and highly complementary strategies how firms could establish their own technology as the standard. The first strategy builds on obtaining the support of other, rival companies and firms producing complementary products, possibly through tacit agreements. The second approach consists of market preemption. Finally, and importantly, the firm aiming at winning the standard war has to make the impression from the very outset that its standard will be accepted at the end of the standard war. Put it another way, marketing and other communication techniques can be employed so as to create self-fulfilling expectations among customers and rivals.

However, different firms may have different strategic windows. These are defined as the time period during which they are able to seize market opportunities in line with their resources and capabilities. Smaller firms tend to have shorter strategic windows because they cannot wait for a long time to introduce new innovations and procure competitive advantage as opposed to large corporations that do not tend to rush in introducing new technologies and take unnecessary risks given their solid financial background.

References

Aghion, P. & Giffith, R. (2005). Competition and Growth, Cambridge, Ma.: MIT Press.

Aghion, P. & Howitt, P. (1992) "A model of growth through creative destruction," *Econometrica, 60,* 323-51.

Aghion, P. & Howitt, P. (1998). Endogenous Growth Theory, Cambridge, Ma.: MIT Press.

Aghion, P. & Howitt, P. (2008). The Economics of Growth, Cambridge, Ma: MIT Press.

Arrow, K. J. (1962). "Economic Welfare and the Allocation of Resources for Invention," in the Rate and Direction of Inventive Activity:Economic and Social Factors, Princeton, N. J.: Princeton Univ. Press, 609-625.

Arrow, K. J. (1962). "Economic Welfare and the Allocation of Resources for Invention," in The Rate and Direction of Inventive Activity Princeton, NJ: Princeton University Press.

Barro, R. & Sala-i-Martin, X. (1995). Economic Growth, New York: McGraw Hill.

Baumol, W. J. (1993). Entrepreneurship, Management and the Structure of Payoffs, Cambridge, Ma.: MIT Press.

Baumol, W. J. (2002). The Free-Market Innovation Machine, Princeton: Princeton Univ. Press.

Dasgupta, P. & Stiglitz, J. E. (1980). "Industrial structure and the nature of innovative activity," *Economic Journal, 90*, 266-293.

Dasgupta, P. & Stiglitz, J. E. (1980). "Industrial structure and the nature of innovative activity," *Economic Journal, 90*, 266-293.

Denicolo, V. & Zanchettin, P. (2006). "Competition, Darwinian Selection and Growth," Working Paper, Univ. of Bologna, Italy.

Etro, F. (2004). "Innovation by Leaders," *Economic Journal, 114*, 281-303.

Foray, D. (2006). The Economics of Knowledge, Cambridge, Ma: MIT Press.

Fudenberg, D. & Tirole, J. (1996). Game Theory, Cambridge, Ma.: MIT Press.

Fudenberg, D., Gilbert, R., Stiglitz, J. & Tirole, J. (1983). "Preemption, leapfrogging and competition in patent races," *European Economic Review, 22,* 3-31.

Gottinger, H. W. (2003). Economies of Network Industries, London: Routledge.

Gottinger, H. W. (2006). Innovation, Technology and Hypercompetition, London: Routledge.

Graham, L. (2013). Lonely Ideas: Can Russia Compete, Cambridge, Ma.: MIT Press.

Grossman, G. M. & Helpman, E. (1993). Innovation and Growth in the Global Economy, Cambridge, Ma: MIT Press.

Grossman, G. M. & Shapiro, C. (1987). "Dynamic R&D competition," *Economic Journal, 97*, 372-387.

Harris, C. & Vickers, J. (1987). "Racing with uncertainty," *Review of Economic Studies 54*, 1-21.

Isaacson, W. (2014). The Innovators, New York: Simon and Schuster.

Jaffe, A. B. & Lerner, J. (2004). Innovation and its Discontents, Princeton: Princeton Univ. Press.

Kamien, M. & Schwartz, N. I. (1982). Market Structure and Innovation, Cambridge: Cambridge Univ. Press.

Katz, M. L. & Shapiro, C. (1987). "R&D Rivalry with Licensing or Imitation," *American Economic Review, 77*, 402-420.

Kirzner, I. M. (1973). Competition and Entrepreneurship, Chicago: Univ. of Chicago Press.

Klemperer, P. (2003). Auctions: Theory and Practice, Princeton, N. J.: Princeton Univ. Press.

Landau, R. & Rosenberg, N. eds. (1986). The Positive Sum Strategy, Washington, D.C.: National Academy Press.

Lucas, R. E. (1988). "On the Mechanics of Economic Development," *Journal of Monetary Economics, 22*(1), 3-42.

Mas-Colell, A., Whinston, M. D. & Green, J. R. (1995). Microeconomic Theory, Oxford: Oxford Univ. Press.

Mokyr, J. (1990). The Lever of Riches, Technological Creativity and Economic Progress, Oxford: Oxford Univ. Press.

Nelson, R. & Winter, S. (1982). An Evolutionary Theory of Economic Change, London: Belknap Press.

North, D. C. (1990). Institutions, institutional change, and economic performance, Cambridge: Cambridge University Press.

North, D. C. (2005). Understanding the Process of Economic Change, Princeton, N. J.: Princeton Univ. Press.

Porter, M. (1990). The Competitive Advantage of Nations, New York: The Free Press.

Reinganum, J. F. (1982). "A Dynamic Game of R and D: Patent Protection and Competitive Behaviour," *Econometrica, 50,* 671-688.

Reinganum, J. F. (1985). "Innovation and industry evolution," *Quarterly Journal of Economics, 99,* 81-99.

Reinganum, J. F. (1989). "The Timing of Innovation," in Schmalensee, R. and R. Willig, eds., *The Handbook of Industrial Organization*, Vol.*1*, Amsterdam: North Holland, Chap.14, 850-908.

Romer, P. M. (1990). "Endogenous Technological Change," *The Journal of Political Economy*, *98*(5), 71-102.

Segerstrom, P. (2006). "Intel Economics" *International Economic Review*, *48* (1), 247–280.

Shapiro, C. & Varian, H. (1999). Information Rules – A Strategic Guide to the Network Economy, Boston: Harvard Business School Press.

Shapiro, C. (1989). "Theories of Oligopoly Behavior," in Schmalensee, R. and R. Willig, eds., *The Handbook of Industrial Organization*, Vol. *1*, Amsterdam: North Holland, Chap.6, 330-414.

Solow, R. M. (1956). "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics, 70*(1), 65-94.

Solow, R. M. (1997). Learning from 'Learning by Doing,' Stanford: Stanford Univ. Press.

Sutton, John. (1998). Technology and Market Structure, Cambridge, Ma.: MIT Press.

Tirole, J. (1988). The Theory of Industrial Organization, Cambridge, Ma.: MIT Press.

Van den Berg, H. (2001). Economic Growth and Development, New York: McGraw Hill.

Vickers, John. (1986). "The evolution of industry structure when there is a sequence of innovations," *Journal of Industrial Economics, 35*, 1-12.

Chapter 6

MECHANISM DESIGN FOR ECONOMIC DEVELOPMENT

There is no economic operating system without an underlying mechanism design.

This chapter shows how and which mechanism design in networks would allow efficient systems to achieve network-centered long-run industrial and economic growth, and why informationally efficient mechanism design of the Hayek-Hurwicz type is a superior vehicle for industry driven economic growth.

What are the major factors for economies to succeed on catching up in growth of GDP or GDP per capita as an indicator of prosperity? As has been broadly covered in the economic growth and development literature, see Barro (1999), Gottinger and Goosen (2012), there have been convergence theories on advanced (developed) economies that have only been partially or conditionally verified for some OECD economies, but equally there have been observations of divergence among some developed and developing (transitional) economies that appear to show a growing gap. Technology adoption within regions of a national economy (say, China) were a major factor of less developed regions to catch up. Catchup metrics could also be identified in disaggregated form next to aggregates such as GDP (total), GDP per head and Human Development Index (HDI). For example, we may list an industry/technology index, an infrastructure indicator (transportation, communication networks), R&D expenditure (government or private industry) and competitiveness levels (global or regional market share of key industries). A process of catching up induced by industrial/technology races may tend to converge over time within a bloc of similar countries if technological and educational endowment is similar. The latter originate in Abramovitz' (1986) famous 'social capabilities' as key industries engage in more incremental and complementary innovation through international trade and foreign direct investment (FDI) spreading to emerging industries in likewise developing economies. In post World War II history what was the mechanism that induced Japan to be on a path of catchup growth in the 1960s and 1970s in terms of per capita income growth? As suggested in the Asian miracle (Krugman, 1994), one factor was input growth in key value-added industries through capital expansion, the other was cumulative technological advancement through largely incremental innovation leading to international competitive advantage in some industries. But as soon as the technological frontier was in close reach it became increasingly difficult to pursue market dominance because of competitive density. Also in a wide array of high-tech markets it became decisive to have integrative or systems technologies to catch complementary and increasing returns markets.

In a historical context, over the past 150 years, a macro catchup process becomes multidimensional, emerging with England, moving to the U.S. and a few European economies, Germany, France, later to Japan and the newly industrialized economies (NIEs) in Asia partly overlapping the BRICs. The argument is that due to expansion of trade links through globalization the catchup process is multidimensional and multispeed. Even for those increasing in overall speed the benchmark of catchup keeps shifting from one or a few to many more. For example, in the past WWII period, industrial and economic growth of Japan was driven toward the US economy less so toward Europe; the pattern of growth and catchup of China hits many more emerging economies with significant potentials. We have modelled this by a more complex differential game as in Gottinger and Goosen (2012, Chapter 2).

As a precursor to economic mechanism design, from an economic history perspective, we first identify A. Gershenkron (1962) who emphasized state action on industrialization setting free guided entrepreneurial activities through targeted industrial policy (along the lines of Japan and later South Korea, Taiwan), see Lee (2013). This compares to Abramovitz'(1986) and Abramovitz and David's (1992) catching up process where self-reinforcing, industry specific, competitive market forces, rather than state action, would initiate and sustain a technological race resulting in leadership positions across a range of industries. Here the roots of growth rest in the microeconomic industry structure and new industry creating high technology entrepreneurship as conceptionally and empirically explored by Scherer (1992, 1999).

The World Bank in the past took a balanced position, propagating the state's role to develop social capital which then by itself creates social capability to induce an upward potential through market forces. A more activistic role – argueing against the 'Washington Consensus - has recently been favored by Justin Lin (2013), the former (Chinese) Chief Economist of the World Bank. Amsden's development statism is a further extension of Gershenkron, catchup viewed as a process of learning how to compete through 'Late Industrialization' (Amsden, 1989). Varieties of the main catchup processes listed here have been reviewed by Burkett and Hart-Landsberg(2003), Fagerberg and Godinho (2004).

In 'Achieving Rapid Growth,' Sachs and Warner (1996), in comparing growing middle income countries, identify major factors such as allocative efficiency (with governmental interaction low, markedly less regulation), high degree of competition, free trade. flexible labor markets, low taxes on labor, capital income and profits. While convergence trends among developing economies may be facilitated by becoming more alike in terms of sources of growth such as technological progress, innovation, levels of physical capital, labor productivity, quality of human capital, extent of trade openess, industrial structure and institutional framework, one clear distinguishing factor may involve increasing returns mechanisms (IRMs). IRMs eliminate any kind of convergence and allow for ongoing quasi-linear growth. Anecdotal observations as put forward by Easterly and Levine (2002) can be summarized as stylized facts on growth mechanisms.

(1) Factor accumulation does not account for growth differences but total factor productivity does for a substantial amount.

(2) There are huge growing differences in GDP per capita. On a global scale divergence and not conditional convergence is the major concern in development policy.
(3) Among developing economies growth is not persistent over time.

If movable through market forces all factors of production flow to the same places suggesting important externalities.

In what follows we look at catchup processes as being primarily induced by technological races establishing new industries that allow a free flow of information, through entrepreneurial activity and innovation. Under these circumstances it is more likely that an economic mechanism design generating more information, more choices, more economic freedom, more market transparency supported by democratic institutions will have a better economic performance record, in a sustainably long run perspective than any other that fails on the information distribution and the strategic incentive side which would coincide with various types of collectivist systems.

6.1. A Simple Mechanism Design for Development and Catchup

In following Campbell (1987), a mechanism design approach for resource allocation mechanisms is a tool to judge the performance of an economic or likewise technological system.

A simple mechanism design as an allocation tool for network economies emanates from the following paradigmatic situation:

Let there be any number of i agents in a network economy that collectively generate demand competing for resources from suppliers. The suppliers announce prices entering a bulletin board accessible to all agents (as a sort of transparent market institution). In a simple form of a trading process we could exhibit a ‘tatonnement process’ on a graph where the agents set up a demand to the suppliers who advertise prices on a Bulletin Board which are converted to new prices in interaction with the agents.

The tatonnement process in economics is a simple form of an algorithmic mechanism design (AMD), Nisan and Ronen (2001), which in computer science (CS) emerges as an offspring to algorithmic game theory (Nisan et al., 2007).

Algorithmic ingredients apply to rational and ‘selfish’ agents having well defined utility functions repesenting preferences over possible outputs of the algorithm. A payment ingredient motivates the agents. Mechanism Design Theory (MDT) aims to show how privately known preferences of the entire population can be aggregated towards a ‘social choice’ that drives the mechanism of the whole economy:

(i) each agent or group of agents have some some private input represented by its type embedded in public knowledge or resource endowment as the ‘social environment,’
(ii) there is a production function (or output specification) that associates each type $t = t^1, \ldots, t^n$ with a set of socially allowable outputs (productions), $o \varepsilon O$, each agent’s preferences are codified as real-valued utility functions noted $u^i (t^i, o)$

The universal function is specified in linear terms as $u^i = r^i + v^i (t^i, o)$ where $r^i \geq 0$ is the initial endowment (or resource), or the agent's wealth, which the agent attempts to optimize.

If we assume that the mechanism follows a 'revelation principle' that is truthful to the extent that the agents report their real 'type' then truth-telling is the only dominant strategy. This is considered a truthful implementation. The societal objective function is simply the aggregation of all agents valuations. A maximizing 'mechanism design process' (MDP) is called utilitarian if its objective function is the sum of all agents' utilities.

In the context of the internet economy a MDP would enable users of network applications to present their demand profile through 'quality of service'(QoS) demands via utility functions defining the system performance requirements (Gottinger, 2013). The resource allocation process then involves economic actors to perform economic optimization given scheduling policies, load balancing and service provisioning.

Distributed algorithmic mechanism design for internet resource allocation in distributed systems (as strictly decentralized auction type systems) is akin to an equilibrium converging market based economy where selfish agents maximize utility and firms seek to maximize profits and the state keeps an economic order providing basic public goods and public safety. According to Feigenbaum et al. (2007, 364)), in an internet economy

> "... it is often best to treat the computational entities as independent and selfish agents, interested in only optimizing their own outcome. As a category of behavior, selfishness lies between the extremes of automatic obedience and byzantine disruption; selfish agents are unwilling to follow a central planner's instructions, but they do not act arbitrarily. Instead their actions are driven by incentives, i.e., the prospect of good or bad outcomes. The field of mechanism design ... has shown how, by carefully constructing economic mechanisms to provide the proper incentives, one can use selfish behavior to guide the system toward a socially desirable outcome."

A distributed algorithmic mechanism design (DAMD) thus consists of three components: a feasible strategy space at the network nodes for each agent (or 'autonomous system'), an aggregated outcome function computed by the mechanism and a set of multi-agent prescribed strategies induced by the mechanism.

For a DAMD in place an internet economy can be shown computationally and informationally efficient in the sense of Hurwicz (1972), Hurwicz and Reiter (2006), corroborated from a computer science view by Conitzer and Sandholm(2002). Furthermore, for efficient micro-macro management it would satisfy 'moral hazard' and incentive based concerns. In view of 'informational constraints,' e.g., adverse selection, it may also be superior in performance since in a market based decentralized capitalistic system due to keen competition among operating managers we would expect a larger selection of high-type performing managers acting on truthful payoff-relevant information than in any collectivist system.

A distributed algorithmic mechanism design being computationally efficient in a large decentralized internet economy is a powerful paradigm to substantiate claims by F. A. Hayek (1945) that an industrialized economy based on market principles achieves an overall better performance than socialist type economies of similar nature and scale. This paradigm of an internet economy historically puts the socialist planning debate in a new light. Historically, by the proposals of socialist planning proponents, it should be conducted on the basis of computational feasibility and superiority to any market type mechanism (Hayek, 1935; Cottrell and Cockschott, 1993).

A ‘Hayek-Hurwicz’ DAMD extends to a dynamic market-type real economy which also invokes highly desirable properties of incentive structures (Arrow and Hurwicz,1960; Myerson, 2008) and knowledge creation through Schumpeterian entrepreneurship. This suggests that a MDP of this type should be more likely to generate a long-run sustainable growth and development process with greater welfare benefits in comparison to what any alternative type collectivist planning would achieve.

In the sections to follow we will explore growth and development generating structures and factors that are compatible with a Hayek-Hurwicz design scheme for a developing world.

6.2. Industrial and Macro Competition

The striking pattern that emerges from innovative activities of firms is their rivalry for a technological leadership position in situations that are best described as races or rival contests (Chapter 5). A race is an interactive pattern characterized by firms or nations constantly trying to get ahead of their rivals, or trying not to fall too far behind. In high-technology industries, where customers are willing to pay a premium for advanced technology, leadership translates into increasing returns in the market through positive network externalities. Abramovitz (1986), in explaining the catchup hypothesis, lays stress on a country’s social capability in terms of years of education as a proxy for technical competence and its institutions. Competing behavior is also a dynamic story of how technology unfolds in an industry. In contrast to any existing way of looking at the evolution of technology, racing behavior, though in character more “a productivity race than a runner’s race” (Abramovitz and David, 1997), recognizes the fundamental importance of strategic interactions between competing firms. Thus firms take their rivals’ actions into account when formulating their own decisions. The importance of this characterization is at least twofold. At one level, racing behavior has implications for appreciating technology strategy at the level of the individual firm; at the other level, for understanding the impact of policies that aim to spur technological innovation in an industry or country.

On a national scale, simple catchup hypotheses have put emphasis on the great potential of adopting unexploited technology in the early stage and the increase of self-limiting power in the later stage. However, the actual growth path of the technological trajectory of a specific economy may be overwhelmingly constrained by social capability. The capability endogenously changes as states of the economy and technology evolve. The success of economic growth due to diffusion of advanced technology or the possibility of leapfrogging is mainly attributable to how the social capability evolves. In other words, which effects become more influential: growing responsiveness to competition or growing obstacles to it on account of vested interests and established positions?

Observations on industrial patterns in Europe, the United States or Asia point to which type of racing behavior is prevalent in global high- technology industries. The pattern evolving from such conduct could be benchmarked against the frontier pursuit type of the global technological leaders. Another observation relates to policy inferences on market structure, entrepreneurship, innovation activity, industrial policy and regulatory frameworks in promoting and hindering industry frontier races in a global industrial context. Does lagging behind one's closest technological rivals' cause a firm to increase its innovative effort? The term 'race' suggests that no single firm would want to fall too far behind, and that everyone would like to get ahead. If a company tries to innovate more when it is behind than when it is ahead, then 'catchup' behavior will be the dominant effect. Once a firm gets far enough ahead of its rivals, then the latter will step up their efforts to get closer. The leading company will slow down its innovative efforts until its competitors have drawn uncomfortably close or have surpassed it. This process repeats itself every time a firm gets far enough ahead of its rivals. Of course, catchup may only consistently apply to the next rivals but will not impact the leader. This is called 'persistent leadership.' On a national level catchup processes like this may not lead to convergence.

An alternative behavior pattern would correspond to a business increasing its innovative effort if it gets far enough ahead, thus making catchup by the lagging companies increasingly difficult. For any of these businesses there appears to be a clear link to market and industry structure, as termed 'intensity of rivalry.'

We investigated two different kinds of races: one that is a frontier race between itself and the technological leader at any point in time ('frontier- sticking' behavior), or it might try to actually usurp the position of the leader by 'leapfrogging' it. When there are disproportionately large payoffs to being in the technical lead (relative to the payoffs that a firm can realize if it is simply close enough to the technical frontier), then one would expect that leapfrogging behavior would occur more frequently than frontier-sticking behavior. Alternatively, racing toward the frontier creates the reputation of being an innovation leader hoping to maintain and increase market share in the future. All attempts to leapfrog the current technological leader might not be successful since many lagging firms might be attempting to leapfrog the leader simultaneously and the leader might be trying to get further ahead simultaneously. Correspondingly, one could distinguish between attempted leapfrogging and realized leapfrogging.

Among the key issues to be addressed is the apparent inability of technology-oriented corporations to maintain leadership in fields that they pioneered. There is a presumption that firms fail to remain competitive because of agency problems or other suboptimal managerial behavior within these organizations. An alternative explanation is that technologically trailing firms, in symmetric competitive situations, will devote greater effort to innovation, so that a failure of technological leaders to maintain their position is an appropriate response to the competitive environment. In asymmetric situations, with entrants challenging incumbents, research does demonstrate that startup firms show a stronger endeavor to close up to or leapfrog the competitors. Such issues highlight the dynamics of the race within the given market structure in any of the areas concerned.

We observe two different kinds of market asymmetries with bearing on racing behavior: risk-driven and resource-based. When the incumbents' profits are large enough and do not vary much with the product characteristics, the entrant is likely to choose the faster option in each stage as long as he has not fallen behind in the contest. In view of resource-based asymmetries, as a firm's stage resource endowment increases, it could use the additional resources to either choose more aggressive targets or to attempt to finish the stage sooner, or both. Previous work has suggested that a firm that surges ahead of its rival increases its investment in R&D and speeds up, while a lagging firm reduces its investment and slows down. Consequently, preceding effort suggests that the lead continues to increase. However, based on related work for the US and Japanese telecommunications industry when duopolistic and monopolistic competition and product system complexity for new products are accounted for, the speeding up of a leading firm occurs only under rare circumstances. For example, a company getting far enough ahead such that the (temporary) monopoly term dominates its payoff expression, will always choose the fast strategy, while a company that gets far enough behind will always choose the aggressive approach. Under these conditions, the lead is likely to continue to increase. If, on the other hand, both monopoly and duopoly profits increase substantially with increased aggressiveness then even large leads can vanish with significant probability.

Overall, this characterization highlights two forces that influence a firm's choices in the various stages: proximity to the finish line and distance between the firms. This probability of reaping monopoly profits is higher the farther ahead a firm is of its rival and even more so the closer the firm is to the finish line. If the lead company is far from the finish line, even a sizeable lead may not translate into the dominance of the monopoly profit term, since there is plenty of time for the lead situation to be reversed, and failure to finish first remains a probable outcome. In contrast, the probability that the lagging company will get to be a monopolist becomes smaller as it falls behind the leader. This raises the following question: what kind of actions cause a firm to get ahead? Intuitively, one would expect that a firm that is ahead of its rival at any time t, in the sense of having completed more stages by time t, is likely to have chosen the faster strategy more often. We will construct numerical estimates of the probability that a leading firm is more likely to have chosen a strategy faster to verify this intuition.

Moving away from the firm-led race patterns revolving in a particular industry to a clustering of racing on an industry level is putting industry in different geo-economics zones against each other and becoming dominant in strategic product/process technologies. Here racing patterns among industries in a relatively free-trade environment could lead to competitive advantages and more wealth creating and accumulating dominance in key product/process technologies in one region at the expense of others. There appears to be a link that individual races on the firm level induce similar races on the industry level and will be a contributing factor to the globalization of network industries.

Thus similar catchup processes are taking place between leaders and followers within a group of industrialized countries in pursuit of higher levels of productivity. Supposing that the level of labour productivity were governed entirely by the level of technology embodied in capital stock, one may consider that the differentials in productivities among countries are caused by the 'technological age' of the stock relative to its 'chronological age.' The technological age of capital is the age of expertise at the time of investment plus years elapsing from that time. Since a leading state may be supposed to be furnished with the capital stock embodying, in each vintage, technology which was 'at the very frontier' at the time of investment, the technological age of the stock is, so to speak, the same as its chronological age.

While a leader is restricted in increasing its productivity by the advance of new technology, trailing countries have the potential to make a larger leap as they are provided with the privilege of exploiting the backlog in addition of the newly developed technology. Hence, followers being behind with a larger gap in technology will have a stronger potential for growth in productivity. The potential, however, will be reduced as the catchup process goes on because the unexploited stock of technology becomes smaller and smaller. However, as new technologies arise and are rapidly adopted in a Schumpeterian process of 'creative destruction,' their network effects induce rapid accelerating and cumulative growth potentials which are catalyzed through industry competition. In the absence of such a process we can explain the tendency to convergence of productivity levels of follower countries. Historically, however, it fails to answer alleged puzzles as to why a country, such as the United States, has preserved the standing of the technological leader for a long time since taking over leadership from Britain in around the end of the nineteenth century and why the shifts have taken place in the ranks of follower states in their relative levels of productivity (i.e., technological gaps between them and the leader). Abramovitz (1986) poses some extensions and qualifications on this simple catchup hypothesis in an attempt to explain these facts. Among other factors than technological backwardness, he lays stress on a country's social capability in terms of years of education as a proxy of technical competence and its political, commercial, industrial, and financial institutions. The social capacity of a state may become stronger or weaker as technological gaps close and thus he argues that the actual catchup process does not provide itself to simple formulation. This view has a common understanding to what another economist, Olson (1996), expresses to be 'public policies and institutions' as his explanation of the great differences in per capita income across countries, stating that any poorer states that implement relatively good economic policies and institutions enjoy rapid catchup growth.

The suggestion should be taken seriously when we wish to understand the technological catching-up to American leadership by Japan, in particular during the post-war period, and explore the possibility of a shift in standing between these two countries. This consideration will directly bear on the future trend of the state of the art which exerts a crucial influence on the development of the world economy (Juma and Clark, 2002; Fagerberg and Godinho, 2004). These explanations notwithstanding, we venture as a major factor for divergent growth processes the level of intensity of the racing process within the most prevalent value-added industries with cross-sectional spillovers. These are the communications and information industries which have been shaped and led by leading American firms and where the rewards benefited their industries and country. Although European and Japanese companies were part of the race they were left behind in core relatively free advanced markets reaping lesser benefits.

For leading emerging economies like India and China with a huge market potential but protected market the race is biased toward national champions.

Steering or guiding the process of racing through the pursuit of industrial policies aiming to increase competitive advantage of respective industries, as having been practiced in Japan, would stimulate catch-up races but appears to be less effective in promoting frontier racing. Another profound reason lies in the phenomenon of network externalities affecting ICT industries. That is, racing ahead of rivals in respective industries may create external economies to the effect that such economies within dominant industries tend to improve their international market position and therefore pull ahead in competitiveness vis-à-vis their (trading) partners.

Racing behavior in leading high-growth network industries by generating frontier positions, create critical cluster and network externalities pipelining through other sectors of the economy and create competitive advantages elsewhere, as supported by the increasing returns debate (Arthur, 1996). In this sense we can speak of positive externalities endogenizing growth of these economies and contributing to competitive advantage. All these characteristics lay the foundations of the 'Network Economy.'

The Network Economy is formed through an ever-emerging and interacting set of increasing returns industries; it is about high-intensity, technology driven-racing, dynamic entrepreneurship, and focused risk-taking through (free) venture capital markets endogenized by societal and institutional support. With the exception of pockets of activity in some parts of Europe (the UK and Scandinavia), and in specific areas such as mobile communications, these ingredients for the Network Economy are only in the early stage of emerging in Continental Europe, and the political mindset in support of the Network Economy is anything but prevalent. As long as we do not see a significant shift toward movements in this direction, Europe will not see the full benefits of the Network Economy within a Global Economy.

Racing behavior on technological positions among firms in high- technology industries, as exemplified by the globally operating telecommunications and computer industries, produce spillover benefits in terms of increasing returns and widespread productivity gains. Due to relentless competition among technological leaders the network effects result in significant advantages in the value added to this industry contributing to faster growth of GDP, and through a flexible labor market, also to employment growth. This constitutes a new paradigm in economic thinking through network economies and is a major gauge to compare the wealth-creating power of the US economy over the past decade against the European and advanced Asian economies. It is interesting to speculate on the implications of the way companies in major high-technology markets, such as telecommunications, split clearly into the two major technology races, with one group of firms clearly lagging the other.

The trajectories of technological evolution certainly seem to suggest that firms from one frontier cannot simply jump to another trajectory. Witness, in this regard, the gradual process necessary for a firm in the catchup race to approach those in the frontier race. There appears to be a frontier 'lock-in,' in that once a company is part of a race, the group of rivals within that same race are the ones whose actions influence that company's strategy the most. Advancing technological capability is a cumulative process. The ability to advance to a given level of technical capability appears to be a function of existing technical capability. Given this path dependence, the question remains: why do some firms apparently choose a path of technological evolution that is less rapid than others? Two sets of possible explanations could be derived from our case analysis, which need not be mutually exclusive. The first explanation lingers primarily on the expensive nature of R&D in industries like telecommunications and computers which rely on novel discovery for their advancement. Firms choosing the catchup race will gain access to a particular technical level later than those choosing the frontier, but will do so at a lower cost.

6.3. TECHNOLOGICAL FRONTIERS

The evolution of a cross section of high technology industries reflects repetitive strategic interactions between companies in a continuous quest to dominate the industry or at least to improve its competitive position through company level and industry level technological evolution. We can observe several racing patterns across industries, each of which is the result of a subset of firms jockeying for a position either as a race leader or for a position near the leader constituting a leadership club. The identification and interpretation of the races relies on the fact that different firms take very different technological paths to target a superior performance level with the reward of increasing market shares, maintaining higher productivity and profitability. In a Schumpeterian framework such races cannot be interpreted in a free-riding situation where one firm expands resources in advancing the state of technology and the others follow closely behind. Such spillover interpretations are suspect when products are in the domain of high complexity, of high risk in succeeding, and different firms typically adopt different procedural and architectural approaches.

The logic underlying this evolution holds in any industry in which two broad sets of conditions are satisfied. First, it pays for a firm to have a technological lead over its rival; it also boosts its market image and enhances its reputational capital. Second, for various levels of technological complexity among the products introduced by various firms, technological complexity can be represented by a multi-criteria performance measure, that is, by a vector-valued distance measure. The collection of performance indicators, parameters, being connected with each other for individual companies form an envelope that shapes a ‘technological frontier.’ The technological frontier is in fact a reasonable indicator of the evolving state of knowledge (technical expertise) in the industry. At any point in time the industry technology frontier (ITF) indicates the degree of technical sophistication of the most advanced products carried by companies in that industry in view of comparable performance standards. Firm level technology frontiers (FTF) are constructed analogously and indicate, at any point in time, the extent of technical sophistication achieved by the firm until that point in time. The evolution of company and industry level frontiers is highly interactive. Groups of company frontiers are seen to co-evolve in a manner that suggests that the respective firms are racing to catch up with, and get ahead of each other.

A data set could focus on a given set of products (systems) by major European, American or Asian enterprises in those industries for a sufficiently representative period of market evolution. In principle, we can identify at least two races in progress in the industries throughout a given period of duration. One comprises the world frontier race in each of those industries, the other, for example, the European frontier race which technically would constitute a subfrontier to the worldwide race. The aggregate technology frontier of the firms in a particular race (that is, ITF) is constructed in a manner similar to the individual FTFs. Essentially, the maximal envelope of the FTFs in a particular race constitutes the ITF for that race. The ITF indicates, as a function of calendar time, the best achievable performance by any firm in the race at a given date.

A statistical profiling of technological evolution and innovation relates to competitive racing among rival companies. Among the (non-exclusive) performance criteria to be assessed are (1) frequency of frontier pushing, (2) technological domination period, (3) innovations vs. imitations in the race, (4) innovation frequency when behind or ahead, (5) nature of jumps, leapfrogging or frontier-sticking, (6) inter-jump times and jump sizes, (7) race closeness measures, (8) interfrontier distance, (9) market leading through ‘market making’ innovations and (10) leadership in ‘innovation markets.’

A race may or may not have different firms in the leadership position at different times. It may be a tighter race at some times than at others, and in general, may exhibit a variety of forms of interesting behavior. While analysis of racing behaviour is left to various interpretations, it is appropriate to ask why the firms are motivated to keep on racing at all. As access to superior technology expands the scope of opportunities available to the firms, the technology can be applied in a range of markets. However, leading edge technology is acquired at a cost. It seems unlikely that all the companies would find it profitable to compete to be at the leading edge all the time. Also not every firm has access to equal capabilities in leveraging a given level of technological resources. Firms may, for example, be expected to differ in their access to complementary assets that allows them to appropriately reap the benefits from their innovation. It is reasonable to assume that whatever the level of competence of a company in exploiting its resources it will be better off the more advanced the technology. Based on this procedure an analysis will show how dynamic competition evolved in the past.

Unlike other (statistical) indicators (such as patent statistics) referring to the degree of competitiveness among industries, regions and countries concerned, the proposed measures cover behavioral dynamic movements in respective industries, and therefore are able to lend intrinsic predictive value to crucial economic variables relating to economic growth and wealth creation. The results are likely to provide strategic support for industrial and technology policy in a regional or national context and will enable policy makers to identify strengths and weaknesses of relevant players and their environments in those markets. While this process looks like a micro representation of dynamic technological evolution driving companies and industries into leadership positions, we may construe an analogous process that drives a region or a nation into advancement on a macro scale in order to achieve a higher level pecking order among its peers. This may allow using the micro foundations of racing as a basis for identifying clubs of nations or regions among them to achieve higher levels and rates of growth.

6.4. Catchup or Leapfrogging

It was Schumpeter (1942) who observed that it is the expectation of supernormal profits from a temporary monopoly position following an innovation that is the chief driver of R & D investment. Along this line, the simplest technology race model can be explained as follows: A number of firms invest in R & D. Their investment results in an innovation with the time spent in R & D subject to some varying level of uncertainty. However, a greater investment reduces the expected time to completion of R & D. The model investigates how many firms will choose to enter such a contest, and how much they will invest. However, despite some extensive theoretical examination of technological races there have been very few empirical studies on this subject (Lerner, 1997) and virtually none in the context of major global industries, and on a comparative basis.

Technological frontiers at the firm and industry race levels offer a powerful tool through which to view evolving technologies within an industry. By providing a benchmarking roadmap that shows where an individual firm is relative to the other firms in the industry, they highlight the importance of strategic interactions in the firm's technology decisions. From the interactive process of racing could emerge various behavioural patterns. Does lagging behind one's closest technological rivals cause a firm to increase its innovative effort? The term 'race' suggests that no single company would want to fall too far behind, and that everyone would like to get ahead. If a firm tries to innovate more when it is behind than when it is ahead, then 'catchup' behavior will be the dominant effect. Once a firm gets ahead of its rivals noticeably, then rivals will step up their efforts to catch up. The leader will slow down its innovative efforts until its rivals have drawn uncomfortably close or have surpassed it. This process repeats itself every time a company gets far enough ahead of its rivals. An alternative behavior pattern would correspond to a firm increasing its innovative effort if it gets far enough ahead, thus making catchup by the lagging firms increasingly difficult. This looks like the ‘Intel Model’ where only the paranoid survive (Grove, 1992). For any of these forms there appears to be a clear link to market and industry structure, as termed ‘intensity of rivalry’ by Kamien and Schwarz (1982).

We group two different kinds of races: one that is a frontier race among leaders and would-be leaders (first league) and another that is a catchup race among laggards and imitators (second league). Though both leagues may play their own game, in a free market contest, it would be possible that a member of the second league may penetrate into the first, as one in the first league may fall back into the second. Another aspect of innovation speed has been addressed by Kessler and Bierly (2002). As a general rule they found that the speed to racing ahead may be less significant the more ‘radical’ (drastic) the innovation appears to be and the more likely it leads to a dominant design. These two forms have been applied empirically to the development of the early Japanese computer industry (Gottinger,1998), that is, a frontier racing model regarding the struggle for technological leadership in the global industry between IBM and ‘Japan Inc.’ guided by MITI (now METI), and a catchup racing model relating to competition among the leading Japanese mainframe manufacturers as laggards.

It is also interesting to distinguish between two sub-categories of catchup behavior. A lagging firm might simply try to close the gap between itself and the technological leader at any point in time (‘frontier-sticking’ behavior), or it might try to actually usurp the position of the leader by ‘leapfrogging’ it. When there are disproportionately large payoffs to being in the technical lead (relative to the payoffs that a firm can realize if it is simply close enough to the technical frontier), then one would expect that leapfrogging behavior would occur more frequently than frontier-sticking behavior (Owen and Ulph, 1994). Alternatively, racing toward the frontier creates the ‘reputation’ of being an innovation leader facilitating to maintain and increase market share in the future (Albach, 1997). All attempts to leapfrog the current technological leader might not be successful since many lagging firms might be attempting to leapfrog the leader simultaneously and the leader might be trying to get further ahead simultaneously. Correspondingly, one should distinguish between attempted leapfrogging and realized leapfrogging. This phenomenon (though dependent on industry structure) appears as the predominant behavior pattern in the US and Japan frontier races (Brezis et al., 1991). Albach (1993) cites studies for Germany that show otherwise.

Leapfrogging behavior influenced by the expected size of payoffs as suggested by Owen and Ulph (1994) might be revised in compliance with the characteristics of industrial structure of the local (regional) markets, the amount of R&D efforts for leapfrogging and the extent of globalization of the industry. Even in the case where the payoffs of being in the technological lead are expected to be disproportionately large, the lagging companies might be satisfied to remain close enough to the leader so as to gain or maintain a share in the local market. This could occur when the amount of R&D efforts (expenditures) required for leapfrogging would be too large for a lagging firm to be viable in the industry and when the local market has not been open enough for global competition: the local market might be protected for the lagging local companies under the auspices of measures of regulation by the government (e.g., government purchasing, controls on foreign capital) and the conditions preferable for these firms (e.g., language, marketing practices).

When the industrial structure is composed of multi-product companies, as for example it used to be in the Japanese computer industry, sub-frontier firms may derive spill over benefits in developing new products in other technologically related fields (e.g., communications equipment, consumer electronic products). These companies may prefer an R&D strategy just to keep up with the technological frontier level through realizing a greater profit stream over a whole range of products.

What are the implications of the way firms split cleanly into the two technology races, with one group clearly lagging the other technologically? The trajectories of technological evolution certainly seem to suggest that firms from one frontier cannot simply jump to another trajectory. Witness, in this regards the gradual process necessary for the companies in the Japanese frontier to catch up with those at the global frontier. There appears to be a front line 'lock-in' in that once a firm is part of a race, the group of rivals within that same race are the ones whose actions influence the firm's strategy the most.

Advancing technological capability is a cumulative process. The ability to advance to a given level of technical capability appears to be a function of existing technical potential. Given this 'path dependence,' the question remains: why do some firms apparently choose a path of technological evolution that is less rapid than others? We propose two sets of possible explanations which need not to be mutually exclusive. The first explanation hinges primarily on the expensive nature of R & D in industries like the computer industry which rely on novel scientific discovery for their advancement. Firms choosing the subfrontier will gain access to a particular technical level later than those choosing the frontier, but will do so at a lower cost. Expending fewer resources on R & D ensures a slower rate of technical evolution. The second explanation relates mainly to technological spillovers. Following the success of the frontier firms in achieving a certain performance level, these become known to the subfrontier firms. In fact, leading edge research in the computer industry is usually reported in patent applications and scientific journals and is widely disseminated throughout the industry. The hypothesis is that partial spillover of knowledge occurs to the subfrontier firms, whose task is then simplified to some extent. Notice that the subfrontier firms still need to race to be technological leaders, as evidenced by the analysis above. This implies that the spillovers are nowhere near perfect. Company specific learning is still the norm. However, it is possible that knowing something about what research avenues have proved successful (for the frontier firms) could greatly ease the task for the firms that follow and try to match the technical level of the frontier company.

6.5. Statistical Metrics of Industrial Racing Patterns

Statistically descriptive measures of racing behavior can be established that reflect the richness of the dynamics of economic growth among competing nations. The point of departure for a statistical analysis of industrial racing patterns is the aggregate technological frontier represented by the national production function as a reasonable indicator of the evolving state of knowledge (technical expertise) in a nation or region which is the weighted aggregate of all industries or activities that themselves are represented by the industry technology frontier (ITF). Firm level technology frontiers (FTF) are constructed analogously and indicate, at any point in time, the weighted contribution of that firm to the industry on standard industry classification.

In this context we define 'race' as a continual contest for technological superiority among nations or regions with key industries. Under this conceptualization a race is characterized by a number of countries whose ITFs remain 'close' together over a period (T) of, say, 25 to 50 years. The distinctive element is that countries engaging in a competition have ITFs substantially closer together than of any country not in the race. A statistical analysis should reflect that a race, as defined, may or may not have different countries in the leadership position at different times. It may be a tighter contest at some times than at others, and in general, may exhibit a variety of forms of industrial behavior. We look for clusters of firms whose ITFs remain close enough throughout the duration (formal measures of closeness are defined and measured). We identify races to take place at any level of industrial performance between the very top and the very bottom throughout 50 years duration that is racing from the bottom to racing to the top.

One comprises the world frontier race in each of those industries, the other a subfrontier race (say, North America, Europe, East Asia, China, India, Latin America, Africa) which technically would constitute a subfrontier to the world, allowing under the best of circumstances for the subfrontier to be the frontier. Since the level and breadth of industrial activity is reflected as an indicator for economic welfare, racing to the top would go parallel with economic growth and welfare enhancing, whereas racing from the bottom would correspond to poverty reduction and avoiding stationary (under)development traps.

6.6. Characterization of Statistical Indicators of Industrial Racing

While a variety of situations are possible, the extremes are the following: (a) one country may push the frontier at all times, with the others following closely behind, (b) some countries share more or less equally in the task of advancing the most value generating industry technology frontiers (ITFs). Depending on the situation the most value generating industries may be high technology based increasing returns or network industries that are able to induce complementary emerging industries with high potentials. Extreme situation (a) corresponds to the existence of a unique technological leader for a particular race, and a number of quick followers. Situation (b), on the other hand, corresponds to the existence of multiple technological leaders.

Assessment of Frontier Pushing: The relevant statistics for the races relate to counting the times the ITFs are pushed forward by countries or regions at large within a global or regional frontier. Frontier pushing can be triggered through industrial policy by governments or well fostered entrepreneurship in an advanced capitalistic system.

Domination Period Statistics: Accepting the view that a country/region has greater potential to earn income and build wealth from its technological position if it is ahead of its race suggests that it would be interesting to examine the duration of time for which a country can expect to remain ahead once it finds itself pushing its ITF. We statistically define the 'domination period' to be the duration of time for which a country leads its particular race. It is interesting to note that the mean domination period is virtually indistinguishable for the three races, and lies between three and four years. A difference of means test cannot reject the hypothesis that the mean years of domination tend to cluster but hardly converge. So countries in each of the races can expect to remain ahead approximately in proportion to their technological capability and more than the amount of time after they have propelled themselves to the front of their respective races. However, the domination period tends to be a more uncertain quantity in the world frontier race, to a lesser degree in the EU frontier race than in any smaller regional races (as evidenced by the lower domination period standard deviation).

*Catchup Statistics***:** If key industries of a country push to innovate more when they are behind than when they are ahead, then catchup behavior will be the dominant effect. For each country/region, these statistics compare the fraction of the total innovations carried out by industries in that country (i.e., the fraction of the total number of times that their ITFs advance) when it was engaging in its race when lagging, with the fraction of times that the country actually led its race. In the absence of catchup behavior, or behavior leading to a country increasingly dominating its rivals, we would expect to see no difference in these fractions. Then the fraction of time that a country is ahead of its race could be an unbiased estimator of the fraction of innovations in its key industries that it engages in when it is ahead. Relevant data, however, suggest that this is usually not the case. They appear to show that the fraction of times a state leads its race at any development level in a group or club is larger than the fraction of innovations that occur when the country is ahead, i.e., more innovations occur when the country is lagging than would be expected in the absence of catch-up or increasing dominance behavior. A major exception would arise if the country would act like an 'Intel Economy,' where unchallenged leadership in key industries creates incentives to increase the lead to its rivals. Catchup behavior is supported by additional observations, as derivable from convergence and conditional convergence in the economic growth process that countries make larger jumps (i.e., the ITFs advance more) when they are behind than when they are leading the race.

*Leapfrogging Statistics***:** From this, the distinction emerges between two kinds of catchup. A lagging country might simply try to close the gap between itself and the technological leader at any point in time (frontier-sticking behavior), or it might try to actually usurp the position of the leader by 'leapfrogging' it. When there are disproportional larger incomes per head when being in the technical lead (relative to a situation that a country can realize if it is simply close enough to the technological frontier), then one would expect that leapfrogging behavior would make it a more attractive incentive than frontier-sticking behavior.

All attempts to leapfrog the current technological leader might not be successful since many lagging firms/industries might be attempting to leapfrog the leader simultaneously. Correspondingly, we observe both the attempted leapfroggings and the realized leapfroggings. It appears likely that the leapfrogging phenomenon would be more predominant in the premier league than in following up leagues.

Interfrontier Distance: How long does 'knowledge' take to spillover from frontier to subfrontier industries? This requires investigating 'interfrontier distance.' One measure of how much subfrontier industries' technology lags the frontier industries' technology could be graphed as 'subfrontier lag' in terms of calendar time. At each point in time, this is simply the absolute difference in the subfrontier performance and the frontier performance time. The graph would clearly indicate that this measure has been declining or increasing more or less monotonically over the past 50 years to the extent that the subfrontier industries have been able/unable to catch up with the frontier industries. A complementary measure would be to assess the difficulty of bridging the lag. That is, how much longer does it take the subfrontier to reach a certain level of technical achievement after the frontier has reached that level? Thus it might very well turn out that the interfrontier distance may be decreasing though the difficulty in bridging the gap is increasing.

Race Closeness Measure (RCM): None of the previous analyses tell us how close any of the overall races are over a period of time. The races are all distant/close by construction, however, some might be closer than others, We define ‘a measure of closeness’ of a race (RCM) at a particular time as follows:

$$\text{RCM}(t) = \Sigma_0^N [F_i(t) - F_j(t)]^2 / N(t) \qquad (1)$$

where F_i (t) is country’s i ITF at time t, $F_j(t)$ is country’s j comparable ITF at time t = max [ITF(t)] for each i, j and N(t) is the number of active key value-generating industries at time t.

The measure (Equation 1) thus constructed has a lowest value of 0, which corresponds to a ‘dead heat’ race. Higher values of the measure correspond to races that are less close. Unlike the earlier characteristics (domination period length, innovation when ahead versus when behind, leapfrogging versus frontier-sticking) which investigate the behavior of a particular feature of the race and of a particular industry in relation to the race frontier, the RCM is more of an aggregate statistic of how close the various racing parties are at a point in time. The closeness measure is simply an indication of parity, and not one that says anything per se about the evolution of the technological frontier. To see this, note that if none of the frontiers were evolving, the closeness measure would be 0, as it would be if all the frontiers were advancing in perfect lock-step with one another.

We talk about value-added increasing returns industries over a period of 30 years. The industries comprise ICT, Consumer Electronics, Chemicals and Materials, Automobiles, Pharma/Biotech, Machine Tools, Medical Instruments, Aerospace/Defense, Energy Technologies, and HT Transportation Systems. Industry sectors can be assigned to various countries/regions such as US, EU, China, Russia, India, Brazil, Japan. We benchmark the industry technology frontiers (ITFs) accordingly, that is, highest ‘state of knowledge’ at time t is 100 pc. The countries’ rank to the max ITFs is assessed as the share of the max ITF. The assessment intervals are spaced in five year intervals starting in 1980 until 2010. After aggregating across industries for each observation point, altogether 7, we get trend graphs over the entire observed period (Table 1).

6.7. Efficiency

Next we explore the inefficiency of the follower nations; i.e., the negative effect on the potential technology gap stemming from inefficient social and institutional factors. A good example of cross industrial inefficiencies over a historically representative period (1810 to 2000) is Russia that was hardly advancing economically against underdeveloped benchmark countries and falling behind leading economies, reinforced through the bolshevik revolution and its underperforming economic mechanism design(Gaidar,2004). Increasing efficiencies deblock catchup in lagging countries (Juma and Clark, 2002). Efficiency is found by dividing a nation's estimated fixed effect by the regional adoption rate. As defined here, it is quite robust to different estimations and samples. The relative efficiencies of the nations within regions appear to conform to common beliefs. For example, in Europe, the Netherlands, Belgium and Switzerland are the most efficient while Turkey, Portugal and Greece are the least efficient. In East Asia, Hong Kong is the most efficient while Indonesia and Thailand are the least efficient. Finally, in Latin America, Mexico and Argentina are at the top and Honduras and Bolivia at the bottom. Another way to discuss the findings is to consider the time required to catchup. Previously, Parente and Prescott (2004) showed that countries with lower levels of relative efficiency will adopt modern technologies at much later dates. Conversely, one could argue that if those countries adopt modern technologies concurrently with their low level of relative efficiency then their rates of growth would stay at a subparity level of their potential.

Table 1. ITF Trends across Industries in Different Industries. The industries comprise ICT, Consumer Electronics, Chemicals and Materials, Automobiles, Pharma/Biotech, Machine Tools, Medical Instruments, Aerospace/Defence, Energy Technologies, HT Transportation Systems

ICT Industry 1980	**ITF (max = 100)**
US	80
EU	60
China	15
USSR	30
India	25
Brazil	20
Japan	70
All 10 Industries	**GDP Share (pc)**
US	70
EU	60
China	50
USSR	40
India	30
Brazil	25
Japan	65

One major source of efficiency generation for a country, according to Parente and Prescott (2004), is belonging to a 'free trade club' that improves efficiency through greater industrial competition. We calculate the required time period until the nations reach their frontier when only the catchup term and inefficiency are allowed to vary across regions and countries. Two frontiers are considered: nations' inefficiency frontier and the leader nation's frontier. The latter requires that the inefficiency levels fade away in time which we assume occurs at the rate of ρ. The European countries, with the exception of Turkey, all seem to have reached their inefficiency reduced frontier. The same is true for most of the East Asian countries. Thus, these nations will not catch up with the U.S. without higher accumulation rates or improved efficiency. For Latin America, most countries are still catching-up with their inefficiency frontier, so that if accumulation rates were the same catchup would still take place through diffusion of disembodied technology. Of course, if inefficiency levels remain then a follower could never completely catch up with the leader by taking advantage of the technology gap alone. As an illustrative example, for the required time to catch-up with the leader if inefficiency levels were improving at the rate ρ much of Europe and Latin America could then approach the frontier faster than East Asia on account of East Asia's lower rate of technology adoption. This begs the question of what determines these (in)efficiencies?

It is reasonable to expect a tradeoff between a general technology level (GTL) of a nation's leading industries and its institutional efficiencies (IE). Thus, using an aggregate score, (GTL, IE), say, a country may be in the top rank of GTL but weak on IE which may be surpassed in growth by one which is lower in GTL rank but strong on IE.

CONCLUSION

Economic growth over the long-run can only be achieved in the course of a real, sustainable value-creating process through industrial performance and open markets in which technology and innovation are the key facilitators. Nations with their industries engage in rival contests in what we term industrial races within a given international trade regime. This reflects a micro-economic based behavioral focus on economic growth (positive or negative). It builds a deeper foundation to explanations of economic growth than purely conventional macroeconomic texts. It also outlines the true sources of growth as a tool for growth diagnostics (Rodrik, 2007, Chapter 2) allowing to embrace other observations on urban growth and non primarily economic factors. In an influential paper in Foreign Affairs entitled 'Can India overtake China' Huang and Khanna (2003) first looked at macro-economic factors, which favor China. They then considered micro-economic structures and behaviors such as competent indigenous entrepreneurship, a sound capital market, an independent legal system, property rights and a grass roots approach to development. The latter all favor India in the long run, say over the next fifty years.

In a widely covered empirical investigation on global growth patterns we concur with Easterly and Levins' (2002) finding that it is not factor accumulation, per se, but total factor productivity that explains cross-country differences in the level of GDP growth rates. This total productivity in turn is derived from technology (innovation) transfer and diffusion, its' supporting institutional characteristics and cultural dependence. Of course, on a deeper level, considerations of merely formal institutions may not suffice for explanations but instead forms of economic mechanism design may be called for that effectively deal with (enforce rules on) 'moral hazard' and 'adverse selection' issues (Myerson, 2008). Economic growth in a decentralized system would be in principle supported by a a Hayek-Hurwicz mechanism design though more in a policy form which is closer to the 'Washington Consensus' (Rodrik, 2007, Chapter 1).

Observations on firm-led racing patterns emerging in oligopolistic market structures of selected high tech industries, and the clustering of racing on an industry level are putting companies in different geo-economic zones against each other, becoming dominant in strategic product/process technologies. Here racing patterns among industries in a relatively free trade environment could lead to competitive advantages and more wealth creating and accumulating dominance in key product / process technologies in one region at the expense of others. The question is whether individual contests on a firm level induce similar effects on an industry level and if so, what controlling effects may be rendered by regional or multilateral policies on regulatory, trade and investment matters? The point is that racing behavior in leading high technology industries by generating frontier positions create cluster and network externalities pipelining through other sectors of the economy and creating competitive advantages elsewhere, as supported by the ‘increasing returns’ debate. In this sense we can speak of positive externalities endogenizing growth of these economies and contributing to competitive advantage.

We are about to show in the upcoming chapters how technological racing, rivalry and competition instigates a process of innovation, industrial and market evolution and how it extends to larger entities than firms and industries to regions and national economies or economy networks. It will show what drives economic growth and globalization, which industries are most significantly affected and how technological racing results in value generation in increasing returns and network industries. Furthermore, we consider how the emergence of selective managerial strategies is most likely to carry success in the pursuit of corporate and industrial policies.

Welfare enhancing technology racing as a constituent element of the capitalist process reinforced by globalization provides social benefits far exceeding the costs. Even more important, any alternative path, other than the competitive, would likely be inferior given the costs in that it would generate a less valued and less welfare producing technology portfolio. That is, even if the competitive process is wasteful, (for example, in parallel or correlated technology development) its unique high value innovation outcome far exceeds the benefits of any alternative path. There is historical, observational and analytical evidence given in Gottinger and Goosen (2012, Chapters, 2-4).

On a national scale simple catch-up hypotheses have put emphasis on the great potential of adopting unexploited technology in the early stage and the increase of self-limiting power in the later stage. However, an actual growth path of technological trajectory of a specific economy may overwhelmingly be constrained by social capability. The capability also endogenously changes as states of the economy and technology evolve. The success of economic growth due to diffusion of advanced technology or the possibility of leapfrogging is mainly attributable to how the social capability evolves (i.e., which effects become more influential: growing responsiveness to competition or growing obstacles to it on account of vested interests and established positions). Another observation relates to policy inferences on market structure, entrepreneurship, innovation activity, industrial policy and regulatory frameworks in promoting and hindering industry frontier races in a global industrial context. Does lagging behind one’s closest technological rivals cause an industry to increase its innovative effort?

On an industry level, among the key issues to be addressed is the apparent inability of technology oriented corporations to maintain leadership in fields that they pioneered. There is a presumption that firms fail to remain competitive because of agency problems or other suboptimal managerial behavior within these organizations. An alternative explanation is that technologically trailing firms, in symmetric competitive situations, will devote greater effort to innovation, so that a failure of technological leaders to maintain their position is an appropriate response to the competitive environment. In asymmetric situations, with entrants challenging incumbents, research does demonstrate that start-up firms show a stronger endeavour to close up to or leapfrog the competitors. Such issues highlight the dynamics of the race within the given market structure in any of the areas concerned.

Catchup processes are taking place between leaders and followers within a group of industrialized countries in pursuit of higher levels of productivity and economic growth. Supposing that the level of labour productivity were governed entirely by the level of technology embodied in capital stock, one may consider that the differentials in productivities among countries are caused by the 'technological age' of the stock used by a country relative to its 'chronological age.' The technological age of capital is a period of technology at the time of investment plus years elapsing from that time. Since a leading country may be supposed to be furnished with the capital stock embodying, in each vintage, technology which was 'at the very frontier' at the time of investment, 'the technological age of the stock is, so to speak, the same as its chronological age.' While a leader is restricted in increasing its productivity by the advance of new technology, trailing countries 'have the potential to make a larger leap' as they are provided with the privilege of exploiting the backlog in addition of the newly developed technology. Hence, followers being behind with a larger gap in technology will have a stronger potential for growth in productivity. The potential, however, will be reduced as the catch-up process goes on because the unexploited stock of technology becomes smaller and smaller. However, as new technologies arise and are rapidly adopted in a Schumpeterian process of 'creative destruction,' their network effects induce rapid accelerating and cumulative growth potentials being catalyzed through industry racing.

In the absence of such a process we can explain the tendency to convergence of productivity levels of follower countries. Historically, it fails to answer alleged puzzles of why a country, such as the United States, has preserved the standing of the technological leader for a long time since taking over leadership from Britain in around the end of the nineteenth century and why the shifts have taken place in the ranks of follower countries in their relative levels of productivity (i.e., technological gaps between them and the leader). Abramovitz (1986) poses some extensions and qualifications on this simple catch-up hypothesis in the attempt to explain these facts. Among other factors than technological backwardness, he lays stress on a country's 'social capability' (i.e., years of education as a proxy of technical competence and its political, commercial, industrial, and financial institutions). The social capability of a country may become stronger or weaker as technological gaps close and thus, he states, the actual catchup process does not lend itself to simple formulation. This view has a common understanding to what another economist, Olson (1996), expresses to be public policies and institutions as his explanation of the great differences in per capita income across countries, stating that any poorer countries that adopt relatively good economic policies and institutions enjoy rapid catchup growth. The suggestion should be taken seriously when we wish to understand the technological catching-up to American leadership by Japan, in particular, during the post-war period and explore the possibility of a shift in standing between these two countries. This consideration will directly bear on the future trend of the state of the art which exerts a crucial influence on the development of the world economy.

These explanations notwithstanding, we venture as a major factor for divergent growth processes the level of intensity of the racing process within the most prevalent value-added industries with cross sectional spillovers. These are the communications and information industries which have been shaped and led by leading American firms and where the rewards benefited their industries and country. Though European and Japanese companies were part of the race they were left behind in core markets reaping lesser benefits. The IT investment relative to GDP, for example, used to be only less than half in countries such as Japan, Germany and France compared to the US. This does not bode well for a rapid catchup in those countries. Steering or guiding the process of racing through the pursuit of industrial policies aiming to increase competitive advantage of respective industries, as having been practised in Japan, would stimulate catch-up races but appears to be less effective in promoting frontier racing. Another profound reason lies in the phenomenon of network externalities affecting IT industries. That is, racing ahead of rivals in respective industries may create external economies to the effect that such economies within dominant industries tend to improve their international market position and therefore pull ahead in competitiveness vis-a-vis their (trading) partners.

The point is that racing behavior in leading high growth network industries by generating frontier positions create critical cluster and network externalities pipelining through other sectors of the economy and creating competitive advantages elsewhere, as supported by the increasing returns debate (Arthur, 1996). In this sense we can speak of positive externalities endogenizing growth of these economies and contributing to competitive advantage.

All these characteristics lay the foundations of the ‘Network Economy.’ The latter is formed through an ever emerging and interacting set of increasing returns industries, it is about high-intensity, technology driven racing, dynamic entrepreneurship, focussed risk-taking through (free) venture capital markets endogenized by societal and institutional support. With the exception of pockets of activity in some parts of Europe (the UK and Scandinavia), and in specific areas such as mobile communication, these ingredients for the Network Economy are only in the early to mid stage of emerging in Continental Europe, and the political mindset in support of the Network Economy is anything but prevalent. As long as we do not see a significant shift toward movements in this direction Europe will not see the full benefits of the Network Economy within a Global Economy.

Racing behavior on technological positions among firms in high technology industries, as exemplified by the globally operating telecommunications, and computer industries, produce spillover benefits in terms of increasing returns and widespread productivity gains. Due to relentless competition among technological leaders the network effects lead to significant advantages in the value added to this industry, contributing to faster growth of GDP, and through a flexible labour market, also to employment growth. This constitutes a new paradigm in economic thinking through network economies and is a major gauge to compare the wealth creating power of the US economy against the European and advanced Asian economies.

REFERENCES

Abramovitz, M. & David, P. (1995). “Convergence and Deferred Catchup,” Chapter 1 in Growth and Development: The Economics of the 21st Century, Ralph Landau, Timothy Taylor, and Gavin Wright, eds., Stanford Ca: Stanford University Press.

Abramovitz, M. (1986). “Catching up, forging ahead, and falling behind,” *Journal of Economic History*, *66*, 385–406.

Albach, H. (1994). Information, Zeit und Wettbewerb (Information, Time and Competition), J. H. von Thuenen Lecture, in Jahrestagung des Vereins fuer Socialpolitik, Muenster, Berlin: Duncker und Humblot, 113–154.

Albach, H. (1997). "Global Competition Among the Few," The Ehrenrooth Lectures, Swedish School of Economics and Business Administration, Res. Reports 40. Helsingfors.

Amsden, A. H. (1989). Asia's New Giant, New York: Oxford Univ. Press.

Arrow, K. J. & Hurwicz, L. (1960). "Decentralization and Computation in Resource Allocation," in Essays in Economics and Econometrics, A Volume in Honor of Harold Hotelling, Chapel Hill: The Univ. of North Carolina Press, 34-104.

Arthur, W. B. (1994). Increasing Returns and Path Dependence in the Economy, Ann Arbor: The Univ. of Michigan Press.

Arthur, W. B. (1996). "Increasing returns and the new world of business," *Harvard Business Review*, July–August, 100–109.

Barro, R. J. (1999). Determinants of Economic Growth, Cambridge: MIT Press.

Brezis, E., Krugman, P. & Tsiddon, D. (1991). "Leapfrogging: a theory of cycles in national technological leadership," *National Bureau of Economic Research (NBER)*, Working Paper # 3886.

Burkett, P. & Hart-Landsberg, M. (2003). "A Critique of 'Catch-Up' Theories of Development," *Journal of Contemporary Asia*, *33*(2), 147-171.

Campbell, D. E. (1987). Resource Allocation Mechanisms, Cambridge: Cambridge Univ. Press.

Conitzer, V. & Sandholm, T. (2002). "Complexity of Mechanism Design," Computer Science Dept., Carnegie Mellon Univ., Pittsburgh, Pa., Disc. Paper, May 28, 1-8.

Cottrell, A. & Cockshott, W. P. (1993). "Calculation, Complexity and Planning: The Socialist Calculation Debate Once Again," *Review of Political Economy, 5*(1), 73-112.

Easterly, W. & Levine, R. (2002). "It's not Factor Accumulation: Stylized Facts and Growth Models," Central Bank of Chile, Working Papers No. 164.

Fagerberg, J. & Godinho, M. M. (2004). "Innovation and Catching-Up," Chap. 20 in Handbook of Innovation, Oxford: Blackwell.

Feigenbaum, J., Schapiro, M. & Shenker, S. (2007). "Distributed Algorithmic Mechanism Design," in Nisan et al. (2007), 363-384.

Gaidar, E. T. (2004). "Modern Economic Growth in Russia and Catch-Up Development," Herald of Europe, Sept.,1-11, also in Gaidar, E. T. (2012), Russia, A Long View, Cambridge: MIT Press.

Gottinger, H. W. & Goosen, M. eds. (2012). Strategies of Economic Growth and Catch-Up, New York: Nova Science.

Gottinger, H. W. (2006a). "Global Technological Races," *Japan and the World Economy, 18,* 181-193.

Gottinger, H. W. (2006b). Innovation, Technology and Hypercompetition, London, Routledge.

Gottinger, H. W. (2013). "Quality of Service on Queueing Networks for the Internet," *iBusiness*, *5*, 1-12.

Grossman, G. & Shapiro, C. (1987). "Dynamic R&D Competition," *Economic Journal*, *97*, 372–387.

Grove, A. S. (1999). Only the Paranoid Survive, New York: Doubleday.

Harris, C. & Vickers, J. (1987). "Racing with Uncertainty," *Review of Economic Studies*, *54*, 305–321.

Hayek, F. A. (1945). "The Use of Knowledge in Society," *American Economic Review*, *35*, 519-530.

Hayek, F. A. ed. (1935). Collectivist Economic Planning, London: Routledge.

Hurwicz, & Reiter, S. (2006). Designing Economic Mechanisms, Cambridge: Cambridge University Press.

Hurwicz, L. (1972). "On Informationally Decentralized Systems," in McGuire, C. B. and R. Radner, Decision and Organization, Amsterdam: North Holland, 297-336.

Juma, C. & Clark, N. (2002). "Technology Catch-Up: Opportunities and Challenges for Developing Countries," Supra Occ. Papers, *Research Centre for the Social Sciences*, Univ. of Edinburg, UK.

Kamien, M. I. & Schwarz, N. L. (1982). Market Structure and Innovation, Cambridge: Cambridge University Press.

Kessler, E. H. & Bierly, P. E. (2002). "Is faster really better? An Empirical Test of the Implications of Innovation Speed," *IEEE Transactions on Engineering Management*, *49*, 2–12.

Krugman, P. (1994). "The Myth of Asia's Miracle," *Foreign Affairs 73*, 63-78.

Lee, K. (2013). Schumpeterian Analysis of Economic Catchup, Cambridge: Cambridge University Press.

Lerner, J. (1997). "An Empirical Exploration of a Technology Race," *The Rand Journal of Economics*, *28*(2), 228–224.

Lin, Y. (2013). Against the Consensus, Cambridge: Cambridge Univ. Press.

Myerson, R. B. (2008). "Perspectives of Mechanism Design in Economic Theory," *American Economic Review*, *98*(3), 586-603.

Nisan, N. & Ronen, A. (2001). "Algorithmic Mechanism Design," *Games and Economic Behavior*, *35*, 166-196.

Nisan, N., Roughgarden, T., Tardos, E. & Vazirani, V. V. eds. (2007). Algorithmic Game Theory, Cambridge: Cambridge University Press.

Olsen, M. (1996). "Big Bills left on the Sidewalk: Why some nations are rich and others poor," *Journal of Economic Perspectives, 10*, 3-24.

Owen, R. & Ulph, D. (1994). "Racing in Two Dimensions," *Journal of Evolutionary Economics*, *4*, 185–206.

Parente, S. L. & Prescott, E. C. (2004). "What a Country Must Do to Catch-Up to the Industrial Leaders," mimeo, Federal Reserve Bank of Minneapolis.

Rodrik, D. (2007). One Economics, Many Recipes: Globalization, Institutions and Economic Growth, Princeton: Princeton Univ. Press.

Sachs, J. D. & Warner, A. M. (1996). "Achieving Rapid Growth in the Transition Economies of Central Europe," *Harvard Institute for International Development, Discussion Paper 544*, July.

Scherer, F. M. (1992). International High-Technology Competition, Cambridge: Harvard University Press.

Scherer, F. M. (1999). *New Perspectives on Economic Growth and Technological Innovation,* Washington D.C.: Brookings Institution Press.

Schumpeter, J. (1942). Capitalism, Socialism and Democracy, New York: Harper, Rev. Edition.

Chapter 7

NETWORK-CENTERED INDUSTRIAL GROWTH

7.1. NETWORK DYNAMICS

We observe that innovation (technology) driven by dynamic competition naturally gives rise to networks in a given economy likely facilitating 'increasing returns' and industrial growth. Conversely, on the formation of networks along this line, constituting some industrial profile, new networks will be building through dynamic competition, and in turn induce changes in industrial profiles. Therefore, network creation as a dynamical system involving an emerging process would result in continuous structural change for an economy and present an engine for industrial growth.

In a pertinent statement on network science (Newman et al., 2006, Chapter1) we are told:

> "... recent work on networks ... takes a dynamical systems view according to which the vertices of a graph represent discrete dynamical entities... Thus a network of interacting individuals, or a computer network in which a virus is spreading, not only has topological properties but has dynamical properties as well."

Put into the economics domain, structural changes in the economy and industrial profiles would be reflected by network changes, the degree of industrial development by network size and density. Globalization in the end would be the outcome of world-wide network effects. Network dynamics in an economics context corroborates with innovation driven micro-economic growth. Thus firm based and industrial growth is one of the major engines of macro-economic growth – what Baumol (2002) calls the 'Free-Market Innovation Machine.' As innovation throughout industries feeds back itself as a self-nourishing process so do new networks build on given networks.

High Speed Technological Competition

On a methodological and instrumental level we see networks behaviorally activated through game theory and mechanism design to operate economic systems.

In highly competitive technological industries new challenges and opportunities are arising in the new product development arena. Driven by global markets, global competition, the global dispersion of scientific/engineering talent, and the advent of new information and communication technologies (ICT) a new vision of product development is that of a highly disaggregated distributive process with people and organizations spread throughout the world. At the same time products are becoming increasingly complex requiring numerous engineering decisions to bring them to market. Competitive pressures mean that ‘time to market’ has become a key to new product success. However, at the same time it is important to keep the innovation and quality dimensions of the new product at their best possible level.

In major parts of competitive analysis involving R&D decisions the focus is on major innovations which could create entirely new markets, for example, in studies featuring patent racing between competing firms. In more common competitive situations we observe firms, however, competing by investing in incremental improvements of products. It is an important aspect when innovation is considered to be manifested in product quality, process improvements and in the overall quality culture of an organization. For example, after product launch, incremental improvement of different aspects of product quality, improvements in various business processes and an incremental adoption of a quality culture are quite real-world phenomena. Thus innovation is conceptualized as being manifested in terms of quality (Gottinger, 2003, Chapter 7). Moreover, the quality dimension extends to pursue continuous total quality improvement for the entire product life cycle. Some firms operate in a simultaneous product launch situation while others compete sequentially by adopting the role of leader or follower.

Over the years with the emergence of e-business and a supply chain view for product development processes, multiple firms with varying and at times conflicting objectives enter into collaborative arrangements. In such situations, the competitive strategy based on quality and innovation could potentially permeate in those collaborative setups. A past example was the collaborative venture of Sony and Samsung to build a cutting edge plant for LCD flat screen TV though displaying ongoing fierce rivalry in new product launches in exactly the same product categories. When innovation and quality levels form the core of a firm’s capabilities, each member in the supply chain would have an incentive to invest and improve their dynamic capabilities. This leads to tacit competition among collaborative product development partners by means of active investment in innovation and quality.

Although the forces of innovation are central to competition in young, technically dynamic industries, they also affect mature industries where life cycles historically were relatively strong, technologies mature, and demands stable.

A strategy for technology must confront primarily what the focus of technical development will be. The question is what technologies are critical to the firm’s competitive advantage. In this context, technology must include the know-how the firm needs to create, produce and market its products and deliver them to customers. As a major step in creating a technology strategy it has to define those capabilities where the firm seeks to achieve a distinctive advantage relative to competitors. For most firms, there are a large number of important areas of technological know-how but only a handful where the firm will seek to create truly superior capability and value.

Having determined the focus of technical development and the source of capability, the firm must establish the timing and frequency for innovation efforts. Part of the timing issue involves developing technical capabilities, and the rest involves introducing technology into the market. The frequency of implementation and associated risks will depend in part on the nature of the technology and the markets involved, but in part on strategic choice. At the extreme, a firm may adopt a rapid incremental strategy, that is, frequent, small changes in technology that cumulatively lead to continuous performance improvement. The polar opposite might be termed the great leap forward strategy. In this approach, a firm chooses to make infrequent but large-scale changes in technology that substantially advance the state of the art.

Thus in a mainstream scenario, high speed technological competition is likely to be driven by

(a) companies competing in fast changing, diverse networks, and by
(b) companies confronted with ever shortening product/technology life cycles because of multiple interactions between current and emerging technologies and product diversity, or what may be circumscribed as 'combinatorial innovation' (Varian et al., 2004).

To represent the features of leader-follower type in a sequence of technological racing we once considered a class of non-cooperative differential games in which some firms have priority of moves over others (Gottinger, 2007).

The firm that has the right to move first is called the leader and the other competing firm is called the follower. A well-known example of this type of sequential move game is the Stackelberg model of duopoly. In this type of interaction the open-loop Nash equilibrium conditions in a sequential move game can be derived. It would lead to a comparison of the strategies of leader and follower. In some alternative context Dasgupta (1986) presents technological competition as bidding games in representing patent races but those games are not truly dynamic because they do not carry a repetitive structure in a changing rivalry. In a differential game the dynamic change of a firm's conduct at each stage is at stake.

Growth Processes

If we look for growth in a high technology industry we may start to single out growth processes for individual companies in that industry. A natural way for a firm to grow is to engage in organic growth through obtaining larger market shares even in a mature or declining industry. This may be achieved through indigeneous innovation on a broader front: technology, organization, services. Second, strategic alliances on major business projects as well as mergers and acquisitions (M&As) can lead companies to new growth opportunities. For example, many well-known companies in the IT business, from Cisco Systems, Google, Intel, Oracle have been consistently growing through M&As even if some have been left under their growth potential. If acquisitions were driven through innovations and entry into new or complementary fields they were targeted at new opportunities and markets. If innovations in-house or through acquisitions led into new complementary markets the resulting growth potential could be characterized as network centered growth, for example through the 'increasing returns' paradigm (Penrose,1995; Roberts, 2006) or through some other form of market (segment) expansion (Gottinger, 2013).

Market Asymmetries and Incentives

Asymmetries between firms are frequently observed in market competition as for example in the case that one firm has a more advanced technology. If there exists some difference between firms, it will certainly affect the incentives to engage in innovative activity. The asymmetries between those firms will generate different incentives to introduce new technology. Since the outcome of technological competition has an obvious effect upon the firms' market positions, firms need to take account not only of the immediate profits that it brings but also the advantage that it may confer in subsequent competition when later innovations are expected.

The analysis uncovers two different forces in determining the incentives to innovate. The competitive incentive is defined as the difference between a firm's profits if it innovates and the profits it would make if its rival innovated instead. The profit incentive is determined by calculating the increase in a firm's 'profits if it alone were investing in R & D. Given those forces, we explore how the incentives to innovate depend on the degree of product differentiation or niche building in product markets.

An alternative outcome could be expected with technological competition when firms can share both the cost and the outcome of innovative activity. The majority of models consider the case that there is only one winner in the technological competition. But alternative models develop technological competition in which each firm engaging in R & D can obtain a patent on a cost reducing technology. Since innovative investments are often very expensive and knowledge has a tendency to leak to others, firms may have an incentive to conduct their innovative activities cooperatively.

Cooperation

Cooperation in R & D among competing firms can produce more rapid technological progress. Since the new technology is shared among competitors through cooperative R & D, firms can avoid the duplication of R & D. The cooperation in R & D can also be a solution to the leakage of information. Since knowledge is inherently a public good, the research done by one firm can be used by another firm even though the latter does not have permission to use the inventive output. Cooperation in R & D is a solution to this problem because cooperative research agreements can internalize spillover externalities. But cooperation between firms' R & D activities may not always produce a better social outcome. The potential gain from R & D will be dissipated to consumers if price competition in the product market is intense. This creates the potential collusive reduction in R & D. That means, firms could use a cooperative agreement in R & D as a vehicle to slow the pace of technological innovation.

In exploring the relationship between technological competition and product differentiation we ask how the rate of product substitutability influences technological progress. The number of firms in the market determines the intensity of price competition. But market competition is also affected by the degree of product differentiation. When the products in the market are less substitutable, the participants in the market will experience less competition. Instead of asking how the incentive to innovate is affected by the increase in the number of firms in the market, which is a common theory in analyzing the relation between R&D and product differentiation (Suzumura, 1995), we consider the question of how the incentive to innovate is related to the degree of product differentiation.

When two products are close substitutes, technological competition is intensified and it drives the market 'to be monopolized' by a firm. If two products are less substitutable, the firms have greater monopolistic power over customers. This discourages a firm from stealing its rival firm's customers through innovation and the market keeps its competitive structure unchanged. It also turns out that the effect of a spillover externality on technological progress is dependent on the closeness of the two products. As stated earlier, knowledge is inherently a public good and it is difficult to prevent its use by others. The analysis shows that the incentives to innovate are generally reduced as the rate of such leakage of knowledge increases. However, when products are very different, the firm that has less of an incentive to innovate has a tendency to raise its R & D effort as the spillover rate increases. The increase in the spillover rate helps the firm that has less of an incentive to improve its marginal contribution of technological progress and compensates for the reduction in its own competence.

Replacement Effect

In the competitive environment of industrial economics between leader and follower, or incumbent and entrant, we observe two contradistinctional behavioural rules. The incumbent's behaviour is influenced by what the literature identifies as the 'replacement effect' (Tirole, 1988, Chap. 10). The conventional replacement effect says that in an effort to maximize the discounted value of its existing profit stream an incumbent monopolist invests less in R&D than an entrant, and thus expects to be replaced by the entrant (for example, when the innovation is drastic enough that the firm with the older technology would not find it profitable to compete with the newer technology). This replacement effect could cause the incumbent to be replaced only temporarily, subsequently he regains a dominant position in the market since he has a superior version of the new technology. The analog event may happen on the macro scale, eventually somewhat more slowly, when one country passes another in innovation induced growth performance. This would be a natural thing consistent with the convergence hypotheses in neoclassical growth models. On the other hand, in the micro industrial economics literature it has been shown that the monopoly term is increasingly important to a firm as it gets ahead of its rival, and that the duopoly term is increasingly important to a firm that falls behind.

Getting Ahead or Catching Up

Another question of interest is whether chance leads to a greater likelihood of increasing the lead, or in more catchup behavior. The existing literature (Grossman and Shapiro, 1987; Harris and Vickers, 1987) has suggested that a firm that surges ahead of its rival increases its investment in R&D and speeds up while a lagging firm reduces its investment and slows down. We call this the Grove paradigm, after Andy Grove (1996) for an Intel economy. On a macro scale an Intel economy would be the prototype of an R&D driven endogeneous growth model when both industry leaders and follower firms invest in R&D in each industry

(Segerstrom, 2006). This behavioral pattern would suggest that the lead continues to increase, and there will be divergence.

On a macro scale this racing paradigm would suggest that the U.S. would mobilize all its technology/science/entrepreneurial resources to increase or keep its distance to other large follow-up economies (such as China or India). With science/ technology being an evolutionary cumulative enterprise, for a dominating country with a portfolio of major increasing returns industries, the odds of leapfrogging oneself are higher than being leapfrogged by close followers, thus this asymmetry could play a distinctive role. Abramovitz' advantage of backwardness may hold on up to a certain limit but with decreasing returns.

Increasing Returns

The Network Economy is formed through an ever-emerging and interacting set of increasing returns industries; it is about high-intensity, technology driven racing, dynamic entrepreneurship, and focused risk-taking through (free) venture capital markets endogenized by societal and institutional support.

Racing behavior on technological positions among firms in high- technology industries, as exemplified by the globally operating telecommunications, computer industries but also biotech industries, produce spillover benefits in terms of increasing returns and widespread productivity gains. Due to relentless competition among technological leaders the network effects result in significant advantages in the value added to this industry, as value networks, contributing to faster growth of GDP, and through a flexible labor market, also to employment growth. This constitutes a new paradigm in economic thinking through network economies and is a major gauge to compare the wealth-creating power of major economic regions in the global context.

The trajectories of technological evolution certainly seem to suggest that firms from one frontier cannot simply jump to another trajectory. Witness, in this regard, the gradual process necessary for a firm in the catchup race to approach those in the frontier race. There appears to be a frontier 'lock-in,' in that once a company is part of a race, the group of rivals within that same race are the ones whose actions influence that company's strategy the most. Advancing technological capability is a cumulative process. The ability to advance to a given level of technical capability appears to be a function of existing technical capability. Given this path dependence, the question remains: why do some firms apparently choose a path of technological evolution that is less rapid than others? Two sets of possible explanations could be derived from our case analysis, which need not be mutually exclusive. The first explanation lingers primarily on the expensive nature of R&D in industries like telecommunications and computers which rely on novel discovery for their advancement. Firms choosing the catchup race will gain access to a particular technical level later than those choosing the frontier, but will do so at a lower cost.

The evolution of a cross section of high technology industries reflects repetitive strategic interactions between companies in a continuous quest to dominate the industry or at least to improve its competitive position through company level and industry level technological evolution. We can observe several racing patterns across industries, each of which is the result of a subset of firms jockeying for a position either as a race leader or for a position near the leader constituting a leadership club. The identification and interpretation of the races relies on the fact that different firms take very different technological paths to target a superior performance level with the reward of increasing market shares, maintaining higher productivity and profitability. In a Schumpeterian framework such races cannot be interpreted in a free-riding situation where one firm expands resources in advancing the state of technology and the others follow closely behind. Such spillover interpretations are suspect when products are in the domain of high complexity, of high risk in succeeding, and different firms typically adopt different procedural and architectural approaches.

Industrial Performance Metrics

The logic underlying this evolution holds in any industry in which two broad sets of conditions are satisfied. First, it pays for a firm to have a technological lead over its rival; it also boosts its market image and enhances its reputational capital. Second, for various levels of technological complexity among the products introduced by various firms, technological complexity can be represented by a multi-criteria performance measure, that is, by a vector-valued distance measure. The collection of performance indicators, parameters, being connected with each other for individual companies form an envelope that shapes a 'technology frontier.' In reference to Chapter 6 the technology frontier is in fact a reasonable statistical metric of the evolving state of knowledge (technical expertise) in the industry. At any point in time the industry technology frontier (ITF) indicates the degree of technical sophistication of the most advanced products carried by companies in that industry in view of comparable performance standards. Firm level technology frontiers (FTF) are constructed analogously and indicate, at any point in time, the extent of technical sophistication achieved by the firm until that point in time. The evolution of company and industry level frontiers is highly interactive. Groups of company frontiers are seen to co-evolve in a manner that suggests that the respective firms are racing to catch up with, and get ahead of each other.

A data set could focus on a given set of products (systems) by major globally operating enterprises in those industries for a sufficiently representative period of market evolution. In principle, we can identify at least two races in progress in the industries throughout a given period of duration. One comprises the world frontier race in each of those industries, the other, for example, the European frontier race which technically would constitute a subfrontier to the worldwide race. The aggregate technology frontier of the firms in a particular race (that is, ITF) is constructed in a manner similar to the individual FTFs. Essentially, the maximal envelope of the FTFs in a particular race constitutes the ITF for that race. The ITF indicates, as a function of calendar time, the best achievable performance by any firm in the race at a given date.

As has been shown by Gottinger and Goosen (2012, Chapter 1) in more detail, a statistical profiling of technological evolution and innovation relates to competitive racing among rival companies. Among the (non-exclusive) performance criteria to be assessed are (1) frequency of frontier pushing, (2) technological domination period, (3) innovations vs. imitations in the race, (4) innovation frequency when behind or ahead, (5) nature of jumps, leapfrogging or frontier-sticking, (6) inter-jump times and jump sizes, (7) race closeness measures, (8) inter-frontier distance, (9) market leading through 'market making' innovations and (10) leadership in 'innovation markets.'

A race may or may not have different firms in the leadership position at different times. It may be a tighter race at some times than at others, and in general, may exhibit a variety of forms of interesting behavior. While analysis of racing behavior is left to various interpretations, it is appropriate to ask why the firms are motivated to keep on racing at all. As access to superior technology expands the scope of opportunities available to the firms, the technology can be applied in a range of markets. However, leading edge technology is acquired at a cost. It seems unlikely that all the companies would find it profitable to compete to be at the leading edge all the time. Also not every firm has access to equal capabilities in leveraging a given level of technological resources. Firms may, for example, be expected to differ in their access to complementary assets that allows them to appropriately reap the benefits from their innovation. It is reasonable to assume that whatever the level of competence of a company in exploiting its resources it will be better off the more advanced the technology. Based on this procedure an analysis will show how dynamic competition evolved in the past.

Unlike other (statistical) indicators (such as patent statistics) referring to the degree of competitiveness among industries, regions and countries concerned, the proposed measures cover behavioral dynamic movements in respective industries, and therefore are able to lend intrinsic predictive value to crucial economic variables relating to economic growth and wealth creation. The results are likely to provide strategic support for industrial and technology policy in a regional or national context and will enable policy makers to identify strengths and weaknesses of relevant players and their environments in those markets. While this process looks like a micro representation of dynamic technological evolution driving companies and industries into leadership positions, we may construe an analogous process that drives a region or a nation into advancement on a macro scale in order to achieve a higher level pecking order among its peers. This may allow using the micro foundations of racing as a basis for identifying clubs of nations or regions among them to achieve higher levels and rates of growth.

Advancing technological capability is a cumulative process. The ability to advance to a given level of technical capability appears to be a function of existing technical potential. Given this ‘path dependence,’ the question remains: why do some firms apparently choose a path of technological evolution that is less rapid than others? We propose two sets of possible explanations which need not to be mutually exclusive. The first explanation hinges primarily on the expensive nature of R & D in industries like the computer industry which rely on novel scientific discovery for their advancement. Firms choosing the subfrontier will gain access to a particular technical level later than those choosing the frontier, but will do so at a lower cost. Expending fewer resources on R & D ensures a slower rate of technical evolution. The second explanation relates mainly to technological spillovers. Following the success of the frontier firms in achieving a certain performance level, these become known to the subfrontier firms. In fact, leading edge research in the computer industry is usually reported in patent applications and scientific journals and is widely disseminated throughout the industry. The hypothesis is that partial spillover of knowledge occurs to the subfrontier firms, whose task is then simplified to some extent. Notice that the subfrontier firms still need to race to be technological leaders, as evidenced by the analysis above. This implies that the spillovers are nowhere near perfect. Company specific learning is still the norm. However, it is possible that knowing something about what research avenues have proved successful (for the frontier firms) could greatly ease the task for the firms that follow and try to match the technical level of the frontier company.

Comparative Growth

In an interesting paper, under some seemingly reasonable assumptions on linear technology investment and dynamic equilibrium path of capital accumulation in a neoclassical type model Lau and Wan (1993, 961) obtain the following results which they argue are fully consistent with empirical growth economics: These are:

> "(a) Not all economies converge in growth with each other. (b) Economies with an initial technical capability will converge in growth with the advanced economies. The difference in per capita output grows exponentially, if the developing economy engages only in imitation (and not innovation). (c) With an initial technical capability there is a 'high growth' period, preceded (followed) by a phase of 'trend acceleration' ('trend deceleration'). (d) With an initial technical capability the technology gap widens forever."

The previous analysis indicates that the possibilities opened through competitive industrial racing are far richer and more surprising than they would emerge from single macro-scale models.

Catchup processes are taking place between leaders and followers within a group of industrialized countries in pursuit of higher levels of productivity and economic growth. Supposing that the level of labour productivity were governed entirely by the level of technology embodied in capital stock, one may consider that the differentials in productivities among countries are caused by the 'technological age' of the stock used by a country relative to its 'chronological age' (Abramovitz, 1986). The technological age of capital is a period of technology at the time of investment plus years elapsing from that time. Since a leading country may be supposed to be furnished with the capital stock embodying, in each vintage, technology which was 'at the very frontier' at the time of investment, 'the technological age of the stock is, so to speak, the same as its chronological age.' While a leader is restricted in increasing its productivity by the advance of new technology, trailing countries 'have the potential to make a larger leap' as they are provided with the privilege of exploiting the backlog in addition of the newly developed technology. Hence, followers being behind with a larger gap in technology will have a stronger potential for growth in productivity. The potential, however, will be reduced as the catch-up process goes on because the unexploited stock of technology becomes smaller and smaller. However, as new technologies arise and are rapidly adopted in a Schumpeterian process of 'creative destruction,' their network effects induce rapid accelerating and cumulative growth potentials being catalyzed through industry racing.

These explanations notwithstanding, we venture as a major factor for divergent growth processes the level of intensity of the racing process within the most prevalent value-added industries with cross sectional spillovers. These are the communications and information industries which have been shaped and led by leading American firms and where the rewards benefited their industries and country. Though European and Japanese companies were part of the race they were left behind in core markets reaping lesser benefits. The IT investment relative to GDP, for example, used to be only less than half in countries such as Japan, Germany and France compared to the US. This does not bode well for a rapid catchup in those countries and they are even less likely to leapfrog. Steering or guiding the process of racing through the pursuit of industrial policies aiming to increase competitive advantage of respective industries, as having been practised in Japan, would stimulate catchup races but appears to be less effective in promoting frontier racing. Another profound reason lies in the phenomenon of network externalities affecting IT industries. That is, racing ahead of rivals in respective industries may create external economies to the effect that such economies within dominant industries tend to improve their international market position and therefore pull ahead in competitiveness vis-a-vis their (trading) partners.

The point is that racing behavior in leading high-tech network industries by generating frontier positions create critical cluster and network externalities pipelining through other sectors of the economy and creating competitive advantages elsewhere, as supported by the increasing returns debate (Arthur, 1996). In this sense we can speak of positive externalities endogenizing growth of these economies and contributing to competitive advantage.

All these characteristics lay the foundations of the 'Network Economy.' The latter is formed through an ever emerging and interacting set of increasing returns industries, it is about high-intensity, technology driven racing, dynamic entrepreneurship, focussed risk-taking through (free) venture capital markets endogenized by societal and institutional support.

7.2. ALLIANCE FORMATION

A discussion on network-centered industrial growth would be incomplete at best if it were not finally to cover the growth impacts of cooperative effects within the firms' competitive framework.

The implications of network strategies vary substantially depending on the purposes for which a particular network of firms forms, as well as the purposes for which the network actually operates (which are not always identical). The creation and management of alliance constellations must be understood in light of motivations for their creation.

While transaction cost economics presents a compelling approach to understanding firm boundaries (Milgrom and Roberts, 1992, Chapter 2), managers clearly engage in mergers and acquisitions (M&A), firm growth and divestiture, re-organization and other boundary shifting initiatives for motivations quite distinct from transaction cost minimization. Marketplace competition encourages firms to minimize costs, but it also compels firms in other directions: to acquire firms to pre-empt or respond to competitors, to adjust strategic vision in the face of disruptive technologies, to innovate to create value over the long term, to name a few. Innovation by its nature requires investment in the creation of new knowledge and capabilities, whether these be capabilities new to a particular firm, or new to the marketplace. In the case of knowledge capital creation, such as in R&D partnerships, 'transaction costs' in the form of inter-firm coordination can in many cases be higher than an internally controlled effort; however, the overall value created by the combination of capabilities between firms can outweigh the increased costs of coordination. The common notion of 'transactional value' attempts to reflect the importance of value calculations in understanding inter-organizational arrangements (Williamson,1995).

While value creation and cost reduction reflect complementary notions for the factors that encourage firms to change their boundaries through M&A or inter-organizational arrangements, these approaches do not assist much in elucidating the complexity of motivations influencing organizational decisions. In broad terms of the field of industrial organization, cost and value creation provide appropriate generalizations. For the purposes of applied corporate strategy, cost and value paradigms are by themselves woefully limiting, particularly given an ever-uncertain future.

Economists attempt to reflect the behavior of market actors based on preferences, available data and other decision factors, as well as provide prescriptions for the most effective economic actions as a result of the insights. The study of corporate strategy attempts to decipher why firms organize and act the way they do, how perhaps they should act regarding the development and execution of effective strategies (but often do not), and the outcomes associated with various scenarios. The fact that much of the strategy literature seeks grounding in economics reflects their complementary nature (McAfee, 2002). Implicit in both of these disciplines is the notion of motivation (Milgrom and Roberts, 1992, Chapter 16). Economic behavior, while observable in the marketplace, can only be modeled in the classical sense of optimizing functions (even given bounded rationality) by making assumptions about the motivations of market actors, known as self-interested behavior. Strategy can only be understood by examining the motivations and purposes for which firms act. The primary motivation of any for-profit firm should be some combination of increasing shareholder value and maximizing profits. This observation is too general to provide helpful insights into network strategy; nonetheless, understanding why firms form alliances to achieve higher profits or shareholder value can help elucidate under what general conditions firms might benefit from an alliance as opposed to governing a resource internally or sourcing it from the market.

Examining ‘economic motivations’ for alliances differs to some extent from understanding the broadest possible set of conditions under which alliances exist. While motivations present a form of conditions, they do not include all conditions impacting alliances. Examples of conditions not subsumed by economic motivations include the personal relationships between firm actors, chance events (that might, however, impact the economic motivations for alliances), and the alliance history of a particular firm (Gulati, 1999). In some cases, personal relationships and social networks could be economically analyzed, by modeling factors such as trust and decreased risk, but the point remains that a number of factors exist which are not primarily economic in nature. Since corporate strategy aims to maximize the economic effectiveness of a for-profit firm, then the most pertinent factors impacting alliances should pertain to the economic and market conditions under which alliances add value.

We can examine these conditions by understanding the motivations for alliance formation. Rather, why do firms form alliances, and how must firms build and execute network strategies under these varied conditions?

Network Formation

What are the key building essentials firms might need for network formation? There are potentially limitless reasons, but all rational, optimizing motivations for forming alliances from a firm's perspective' can be categorized into three types, aside from purely financial motivations

1. *Network Economics*: A firm is attempting to compete in some manner under conditions influenced by network economies.
2. *Innovation/Competencies*: A firm is attempting to augment, transform or further leverage its set of internal competencies in some manner.
3. *Market Structure*: A firm is attempting to compete or become involved in some manner in markets where the market structure compels a firm to form alliances.

Most prominently, this category includes industry structure, positioning (Porter, 1980, 1998), and institutional issues that impact competition, such as government regulation and political risk (North, 1990).

The point is that these three categories - Network Economics, Competencies and Market Structure - provide the dimensions that most powerfully present implications for the creation and prosecution of network strategy. Competencies refer to the internal decisions a firm makes regarding the competencies it develops in-house, which competencies it chooses to source through partners, and which competencies it might leave for the market to provide.

Market Structure refers to the environmental, institutional and competitive factors of a particular industry, market and/or economy. Network Economics refers to whether network economic phenomena significantly influence the dynamics of a particular industry or market under consideration. Network economics includes all of issues traditionally associated with this field of economics, such as demand-side economies of scale, network effects, and positive and negative feedback loops (Shapiro and Varian, 1999; Gottinger, 2003). Network economics could be collapsed into the Market Structure dimension, but only at the expense of crucial explanatory power. The presence or absence of strong network economic effects on a market exerts such a strong influence on the formation of firm networks that it requires its own categorization.

Network Economics

Industries heavily impacted by network economics are characterized by demand-side economies of scale, where the relative number of users can significantly impact the success or failure of a venture. Network economics most directly pertains to industries characterized by technologies which benefit from broad based standards and interoperability, and/or are information intensive. Shapiro and Varian's summary of the implications of network economics on information -intensive industries, Information Rules, underscores how strategy in markets influenced by network economics can differ substantially from those without such dynamics.

For instance, in some cases an open, give-it-away strategy can provide superior long-term prospects over a proprietary approach. This distinction presents interesting challenges to the definition and application of 'firm-specific' assets, from the resource-based view of firm strategy.

Sun Microsystem's' JAVA programming environment provided a prototype. Sun had thus far invested hundreds of millions of dollars developing and promulgating JAVA with developers, succeeding in creating wide popularity and a broad installed user base. Could JAVA accurately be considered a firm-specific asset, given that Sun had pursued an open source strategy to encourage proliferation of the language? Despite the fact that the firm still maintained various intellectual property rights to the language, Sun had found some difficulty capitalizing on JAVA's success.

The wide-spread success of JAVA within the developer community arose in large part from the open source strategy pursued by Sun, which allowed easy access to the development environment. Sun had pursued a network flexible strategy that has succeeded in driving diffusion of the technology. While widespread adoption had presented Sun with many opportunities, it had also encountered significant challenges in monetizing its success. Sun's variable success with JAVA had been significantly challenged by Microsoft's attempts to usurp the programming environment by amending its application within the Microsoft environment. Much later, in mid 2001, Microsoft announced that JAVA would no longer be supported on Microsoft's newly introduced XP operating system, a critical challenge to JAVA, given Microsoft's then dominating position.

Market Structure

In many industries, market structure issues encourage firms to create alliance constellations. The monopoly position of Microsoft in the last section provides an example. Significant aggregation of market power, as in monopolies and monopsonies, can create situations in which firms that desire to enter these markets have very little choice but to ally with an existing player. Any government mandated monopoly falls into this category as well, but can also overlap with other categories, as in the case of national telecommunications monopolies (network economics), or national control of raw materials mining and export. In the case of an oligopoly, existing players might, officially or unofficially, ally to maintain the status quo, as in the case of a cartel. The opposite can also be true. In highly disaggregated markets, firms might decide to ally in order to create economies of scale, lowering costs and developing bargaining power against suppliers. Economists use metrics such as the Herfindahl Index or a top four-firm concentration ratio to represent market concentration.

An extensive literature exists on the implications of the level of market concentration in the industrial organization literature (Tirole, 1988). The purpose here is not to develop and test many possible hypotheses related to market concentration and alliance formation, but rather to underscore that market structure plays a role in motivating the creation of networks of firms, which holds implications for network strategy.

Market structure dimensions are not limited to monopolistic or oligopolistic conditions. As evident from examples presented earlier, the 'market structure' category includes institutional influences such as government regulation and anti-trust issues that might motivate inter-organizational arrangements over acquisition or market relationships. Here we define 'institutional' in the terms developed by North in the economic field of institutional analysis (North, 1990). These institutional issues relate to the market structure category of the

Taxonomy given the fact that they influence the organizational decisions between firms (i.e., to integrate or ally) by impacting competitive market conditions, providing rules of the game. In this sense, it is not necessary to separate market structure and institutional issues into separate categories. The insights turn out on close inspection to be similar.

Competencies

The competencies dimension addresses firms allying in order to leverage competencies and knowledge between firms, create and/or learn new competencies and knowledge, or a combination of both motivations. It should be noted that alliances formed for specific knowledge creation purposes, such as to license or share intellectual property fit into this mould. Knowledge or intellectual property does not necessarily refer to a 'competency' per se; nonetheless, competencies represent the broader application of knowledge sets and capabilities to accomplish objectives through operational, managerial and learning processes. As such, the Competencies dimension includes any objectives for alliances characterized by knowledge creation, learning and/or transfer.

Competency motivated firm networks take many forms. The simplest outsourcing arrangements can resemble market relationships, while some complex, long-term cooperative arrangements can resemble single firms. In fact, equity joint ventures are often structured as single firms, where the parent, cooperative firm owns equity interests and contribute capital, competencies and even people. The proliferation of outsourcing over the past few decades has occurred largely as a result of firms' interests in shedding internal resources dedicated to secondary or periphery competencies, such as data storage or call-center management. EDS, a pioneer in information technology outsourcing, and EMC, an early leader in enterprise data storage, illustrate firms competing based on these types of relationships. In other cases, competency-based alliances reflect more than the ¯hands-off¯ character of basic outsourcing relationships. The U.S. automobile industry evolved over the final decades of the twentieth century to include fewer suppliers, longer term relationships and significant supplier involvement in the design process. All of these developments greatly improved the industry's global competitiveness (Gulati et al., 2000). But the industry's re-organization has not been limited to increased integration between firms.

REFERENCES

Abramovitz, M. (1986). 'Catching up, forging ahead, and falling behind,' *Journal of Economic History*, *66*, 385–406.

Baumol, W. J. (2002). The Free-Market Innovation Machine: Analyzing the Growth Miracle of Capitalism, Princeton and Oxford: Princeton Univ. Press.

Dasgupta, P. (1986). "The Theory of Technological Competition," in Binmore, K. and P. Dasgupta (eds.), Economic Organizations as Games, Oxford: Basil Blackwell, 139-164.

Gottinger, H. W. & Goosen, M. F. A. eds. (2012). Strategies of Economic Growth and Catch-Up, New York: Nova Science.

Gottinger, H. W. (2003). Economies of Network Industries, London: Routledge.

Gottinger, H. W. (2007). “High Speed Technology Competition,” *International Journal of Business and Systems*, *10*, 1-16.

Gottinger, H. W. (2013). “Chinese Innovation and Competition: an industry case of the global telecommunications equipment industry,” *International Business and Economics Research, 2*(2), 22-32, doi:10.11648/j.iber.20130202.12.

Grossman, G. & Shapiro, C. (1987). “Dynamic R&D Competition,” *Economic Journal*, *97,* 372-387.

Grove, A. (1996). Only the Paranoid Survive, New York: Doubleday.

Gulati, R. (1999). “Network Location and Learning. The Influence of Network Resources and Firm Capabilities on Alliance Formation,” *Strategic Management Journal, 20*, 397 – 420.

Gulati, R., Nohria, N. & Zaheer, A. (2000). “Strategic Networks,” *Strategic Management Journal, 21*, 203-215.

Harris, C. & Vickers, J. (1987). “Racing with Uncertainty,” *Review of Economic Studies*, *54,* 305-321.

Lau, M. L. & Wan, H. (1993). “On the Mechanism of Catching Up,” European Economic Review 38, 952-963, also Chap.2 in Wan, H. Y. (2006), Harnessing Globalization, Singapore: World Scientific.

McAfee, Preston R. (2002). Competitive Solutions, Princeton: Princeton Univ. Press.

Milgrom, P. & Roberts, J. (1992). Economics, Organization and Management, Englewood Cliffs, NJ: Prentice Hall.

Newman, M., Barabasi, A. L. & Watts, D. J. eds. (2006). The Structure and Dynamics of Networks, Princeton and Oxford: Princeton Univ. Press.

North, D. C. (1990). Institutions, Institutional Change and Economic Performance, Cambridge: Cambridge Univ. Press.

Penrose, E. (1995). The Theory of the Growth of the Firm, Oxford: Oxford Univ. Press.

Porter, M. E. (1980). Competitive Strategy, New York: Free Press.

Porter, M. E. (1998). On Competition, Cambridge, Ma.: A Harvard Business Review Book.

Reinganum, J. (1989). “The Timing of Innovation,” in *Handbook of Industrial Organization*, Vol. *I,* Schmalensee, R. and R. Willig eds., North Holland: Amsterdam, Chap. 14.

Roberts, J. (2006). The Modern Firm, Oxford: Oxford Univ. Press.

Shapiro, C. & Varian, H. R. (1999). Information Rules, Boston: Harvard Business School Press.

Suzumura, K. (1995). Competition, Commitment and Welfare, Oxford: Clarendon Press.

Tirole, J. (1988). The Theory of Industrial Organization, Cambridge, Ma..: MIT Press.

Varian, H. & Shapiro, C. (2004). The Economics of Information Technology, Cambridge: Cambridge Univ. Press.

Williamson, O. E. ed. (1995). Organization Theory, New York: Oxford Univ. Press.

AUTHOR'S CONTACT INFORMATION

Dr. Hans W. Gottinger
Professor of Economics
STRATEC Munich Germany
Email: stratec_c@yahoo.com

INDEX

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R