PART ONE

# Competition and market power

Hal R. Varian

## 1 Introduction

During the 1990s there were three back-to-back events that stimulated investment in information technology: telecommunications deregulation in 1996, the "year 2K" problem in 1998–99, and the "dot com" boom in 1999–2000. The resulting investment boom led to a dramatic run-up of stock prices for information technology companies.

Many IT companies listed their stocks on NASDAQ. Figure 1 depicts the cumulative rate of return on the NASDAQ and the S&P 500 during most of the 1990s. Note how closely the two indices track each other up until January of 1999, at which point NASDAQ took off on its roller-coaster ride. Eventually it came crashing back, but it is interesting to observe that the total return on the two markets over the eight years depicted in the figure ended up being about the same.
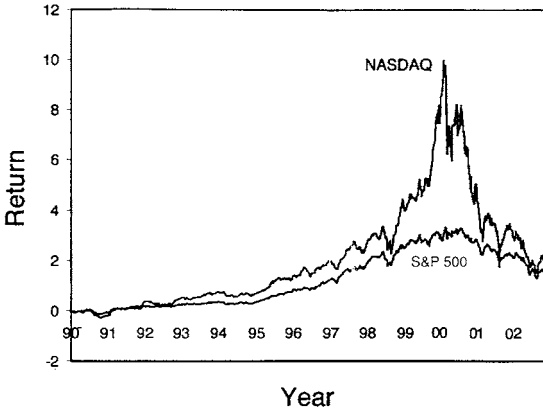
## Cumulative Returns



Figure 1   Return on the NASDAQ and S&P 500 during the 1990s

Figure 1 actually understates the magnitude of technology firms on stock market performance, since a significant part of the S&P return was also driven by technology stocks. In December 1990, the technology component of the S&P was only 6.5 percent; by March 2000, it was over 34 percent. By July 2001, it was about 17 percent.

A prominent Silicon Valley venture capitalist described the dramatic run-up in technology stocks as the "greatest legal creation of wealth in human history." As subsequent events showed, not all of it was legal and not all of it was wealth.

But the fact that only a few companies succeeded in capitalizing on the Internet boom does not mean that there was no social value in the investment that took place during 1999–2001. Indeed, quite the opposite is true. One can interpret figure 1 as showing something quite different from the usual interpretation, namely that competition worked very well during this period, so that much of the social gain from Internet technology ended up being passed along to consumers, leaving little surplus in the hands of investors.

Clearly the world changed dramatically in just a few short years. Email has become the communication tool of choice for many organizations. The World Wide Web, once just a

scientific curiosum, has now become an indispensable tool for information workers. Instant messaging has changed the way our children communicate and is beginning to affect business communication.

Many macroeconomists attribute the increase in productivity growth in the late 1990s to the investment in IT during the first half of that decade. If this is true, then it is very good news, since it suggests we have yet to reap the benefits of the IT investment of the late 1990s.[1]

## 2  Technology and market structure

A major focus of this monograph is the relationship between technology and market structure. High-technology industries are subject to the same market forces as every other industry. However, there are some forces that are particularly important in high-tech, and these will be our primary concern. These forces are not "new"; indeed, the forces at work in network industries in the 1990s are very similar to those that confronted the telephone and wireless industries in the 1890s.

But forces that were relatively minor in the industrial economy turn out to be critical in the information economy. Second-order effects for industrial goods are often first-order effects for information goods.

Take, for example, cost structures. Constant fixed costs and zero marginal costs are common assumptions for textbook analysis, but are rarely observed for physical products since there are capacity constraints in nearly every production process. But for information goods, this sort of cost structure is very common – indeed, it is the baseline case. This is true not just for pure information goods, but even for physical goods such as silicon chips. A chip fabrication plant can cost several billion dollars to construct and outfit, but producing an incremental chip only costs a few dollars.

---

[1] I will not address the literature on productivity in this survey; see Brynjolfsson and Hitt (2000), Steindel and Stiroh (2001), and Stiroh (2001) for an introduction to this literature. For different approaches, see Litan and Rivlin (2001) and Litan and Varian (2001).

4    HAL R. VARIAN

It is rare to find cost structures this extreme outside of technology and information industries.

The effects I will discuss involve pricing, switching costs, scale economies, transactions costs, system coordination, and contracting. Each of these topics has been extensively studied in the economics literature. I do not pretend to offer a complete survey of the relevant literature, but will focus on relatively recent material in order to present a snapshot of the state of the art of research in these areas.

I try to refer to particularly significant contributions and other more comprehensive surveys. The intent is to provide an overview of the issues for an economically literate, but non-specialist, audience.

For a step up in technical complexity, I can recommend the survey of network industries in the *Journal of Economic Literature* consisting of articles by Katz and Shapiro (1994), Besen and Farrell (1994), Leibowitz and Margolis (1990), and the books by Shy (2001) and Vulkan (2003). Farrell and Klemperer (2003) contains a detailed survey of work involving switching costs and network effects with an extensive bibliography.

For a step down in technical complexity, but with much more emphasis on business strategy, I can recommend Shapiro and Varian (1998a), which contains many real-world examples.

### 3  Intellectual property

Information technology is used to manipulate information. Some of that information may be intellectual property. It follows that the terms and conditions of use for intellectual property play a critical role in the economics of information technology.

Copyright law defines the property rights of the product being sold. Patent law defines the conditions that affect the incentives for, and constraints on, innovation in physical devices and, increasingly, in software and business processes.

I do not directly address intellectual property issues here, but my two co-authors, Joseph Farrell and Carl Shapiro do an admirable job in part II. In addition to their contribution, I can refer the reader

to the surveys by Gallini and Scotchmer (2001), Gallini (2002), and Menell (2000), and to the reviews by Shapiro (2000, 2001a). Samuelson and Varian (2002) describe some recent developments in intellectual property policy.

### 4  The Internet boom

First, we must confront the question of what happened during the late 1990s. Viewed from 2003, such an exercise is undoubtedly premature, and must be regarded as somewhat speculative. No doubt a clearer view will emerge as we gain better perspective on the period, but here I will offer one approach to understanding what went on.

I interpret the Internet boom of the late 1990s as an instance of what one might call "combinatorial innovation."

Every now and then a technology, or set of technologies, emerges whose rich set of components can be combined and recombined to create new products. The arrival of these components then sets off a technology boom as innovators work through the possibilities.

This is, of course, an old idea in economic history. Schumpeter (1934, p. 66) refers to "new combinations of productive means." More recently, Weitzman (1998) used the term "recombinant growth." Gilfillan (1935), Usher (1954), Kauffman (1995) and many others describe variations on essentially the same idea. The concept of "General Purpose Technologies" described in Bresnahan and Trajtenberg (1995) and Helpman (1998) is, in our terminology a particularly important type of component for combinatorial innovation.

The attempts to develop interchangeable parts during the early nineteenth century is a good example of a technology revolution driven by combinatorial innovation.[2] The gradual standardization of design of gears, pulleys, chains, cams, and other mechanical devices led to the development of the so-called "American system

---

[2] See Hounshell (1984) for the fascinating history of technological development during this period.

of manufacture" which started in the weapons manufacturing plants of New England but eventually led to a thriving industry in domestic appliances.

A century later the development of the gasoline engine led to another wave of combinatorial innovation as it was incorporated into a variety of devices from motorcycles to automobiles to airplanes.

As Schumpeter points out in several of his writings (e.g. Shumpeter, 2000), combinatorial innovation is one of the important reasons why inventions appear in waves, or "clusters," as he calls them:

> [A]s soon as the various kinds of social resistance to something that is fundamentally new and untried have been overcome, it is much easier not only to do the same thing again but also to do *similar* things in different directions, so that a first success will always produce a cluster. (p. 142)

Schumpeter emphasizes a "demand-side" explanation for such clustering of innovation. One might also consider a complementary "supply-side" explanation: since innovators are, in many cases, working with the same components, it is not surprising to see simultaneous innovation, with several innovators coming up with essentially the same invention at almost the same time. There are many well-known examples, including the electric light, the airplane, the automobile, and the telephone.

A third explanation for waves of innovation involves the development of complements. When automobiles started to become popular in the early 1900s, where did the paved roads and gasoline engines come from? The answer: the roads were initially the result of the prior decade's bicycle boom, and gasoline was often available at the general store to fuel stationary engines used on farms. These complementary products (and others, such as pneumatic tires) were enough to get the nascent technology going; and once the growth in the automobile industry took off it stimulated further demand for roads, gasoline, oil, and other complementary products. This is an example of an "indirect network effect," which I will examine further in section 10.

The steam engine and the electrical engine also ignited rapid periods of combinatorial innovation. In the middle of the twentieth century, the integrated circuit had a huge impact on the electronics industry. Moore's law has driven the development of ever-more-powerful microelectronic devices, revolutionizing both the communications and the computer industry.

The routers that laid the groundwork for the Internet, the servers that dished up information, and the computers that individuals used to access this information were all enabled by the microprocessor.

But all of these technological revolutions took years, sometimes decades, to work themselves out. As Hounshell (1984) documents, interchangeable parts took over a century to become truly reliable. Gasoline engines took decades to develop. The microelectronics industry took thirty years to reach its current position.

But the Internet revolution took only a few years. Why was it so rapid compared to the others? One hypothesis is that the Internet revolution was minor compared to the great technological developments of the past. (See, for example, Gordon, 2000.) This may yet prove to be true – it's hard to tell at this point.

Another explanation is that the component parts of the Internet revolution were quite different from the mechanical or electrical devices that drove previous periods of combinatorial growth. The components of the Internet revolution were not physical devices at all. Instead they were "just bits." They were ideas, standards, specifications, protocols, programming languages, and software.

For such immaterial components there were no delays in manufacture, or shipping costs, or inventory problems. Unlike gears and pulleys, you can never run out of HTML! A new piece of software could be sent around the world in seconds and innovators everywhere could combine and recombine this software with other components to create a host of new applications.

Web pages, chat rooms, clickable images, web mail, MP3 files, online auctions and exchanges, blogs, wikis, . . . the list goes on and on. The important point is that all of these applications were developed from a few basic tools and protocols. They are the result of the combinatorial innovation set off by the Internet, just as the sewing machine was a result of the combinatorial innovation set

off by the push for interchangeable parts in the late-eighteenth-century munitions industry.

Given the lack of physical constraints, it is no wonder that the Internet boom proceeded so rapidly. Indeed, the rapid pace of innovation continues today. As better and more powerful tools for managing and manipulating web sites have been developed, the pace of innovation has even increased, since a broader segment of the population has been able to create online software applications easily and quickly.

Twenty years ago the very idea that a loosely coupled community of programmers, with no centralized direction or authority, could develop an entire operating system would have been rejected out of hand. Such a development would have been just too absurd. But it has happened: the GNU/Linux operating system was not only created online, but has even become respectable and raised a serious threat to very powerful incumbents.

Such open-source software is like the primordial soup for combinatorial innovation. All the components are floating around in the broth, bumping up against each other and creating new molecular structures, which themselves become components for future development.

Unlike closed-source software, open source allows programmers (and "wannabe programmers") to look inside the black box to see how the applications are assembled. Such knowledge is a tremendous spur to education and innovation.

It has always been so. Look at Josephson's description of the methods of Thomas Edison:

> As he worked constantly over such machines, certain original insights came to him; by dint of many trials, materials long known to others, constructions long accepted were put together in a different way – and there you had an invention.
>
> (Josephson, 1959, p. 91)

Open source makes the inner workings of software apparent, allowing future Edisons to build on, improve, and use existing programs – combining them to create novel innovations.

One force that undoubtedly led to the very rapid expansion of the web was the fact that HTML was, by construction, open

source. From conception, web browsers have enabled users to "view source," which meant that many innovations in design or functionality could immediately be adopted by imitators – and innovators – around the globe.

Perl, Python, Ruby, and other interpreted languages have the same characteristic. There is no "binary code" to hide the design of the original author. This allows subsequent users to add on to programs and systems, improving them and making them more powerful.

### 4.1 Financial speculation

Each of the periods of combinatorial innovation referred to in the previous section was accompanied by financial speculation. New technologies that capture the public imagination inevitably lead to an investment boom: sewing machines, the telegraph, the railroad, the automobile . . . the list could be extended indefinitely.

Perhaps the period that bears the most resemblance to the Internet boom is the so-called "Euphoria of 1923," when it was just becoming apparent that broadcast radio could be the next big thing.

The challenge with broadcast radio, as with the Internet, was how to make money from it. *Wireless World*, a hobbyist magazine, even sponsored a contest to determine the best business model for radio. The winning idea was "a tax on vacuum tubes" with radio commercials being one of the more unpopular choices.[3]

Broadcast radio, of course, set off its own stock market bubble. When the public gets excited about a new technology, a lot of "dumb money" comes into the stock market. Bubbles are a common outcome. It may be true that it's hard to start a bubble with rational investors – but not it's not that hard with real people.

Though billions of dollars were lost during the Internet bubble, a substantial fraction of the investment made during this period still has social value. Much has been made of the miles laid of

[3] See Smulyan (1994) for a detailed history and Hanson (1998) for a useful overview of this period.

"dark fiber." But it's just as cheap to lay 128 strands of fiber as a single strand, and the *marginal* cost of the "excess" investment is rather low.

The biggest capital investment during the bubble years was probably in human capital. The rush for financial success led to a whole generation of young adults immersing themselves in technology. Just as it was important for teenagers to know about radio during the 1920s and automobiles in the 1950s, it was important to know about computers during the 1990s. "Being digital" (whatever that meant) was clearly cool in the 1990s, just as "being mechanical" was cool in the 1950s.

This knowledge of, and facility with, computers will have large payoffs in the future. It may well be that part of the surge in productivity observed in the late 1990s came from the human capital invested in facility with spreadsheets and web pages, rather than the physical capital represented by PCs and routers. Since the hardware, the software, and the wetware – the human capital – are inexorably linked, it is almost impossible to subject this hypothesis to an econometric test.

### 4.2  Where are we now?

As we have seen, the confluence of Moore's law, the Internet, digital awareness, and the financial markets led to a period of rapid innovation. The result was excess capacity in virtually every dimension: compute cycles, bandwidth, and even HTML programmers. All of these things are still valuable – they're just not the source of profit that investors once thought, or hoped, that they would be.

We are now in a period of consolidation. These assets have been, and will continue to be, marked to market, to better reflect their true asset value – their potential for future earnings. This process is painful, to be sure, but not that different in principle from what happened to the automobile market or the radio market in the 1930s. We still drive automobiles and listen to the radio, and it is likely that the web – or its successor – will continue to be used in the decades to come.