Standardization in Technology-Based Markets

Gregory Tassey

Senior Economist
National Institute of Standards and Technology
(gtassey@nist.gov)

June 1999

[Forthcoming in *Research Policy*]

Standardization in Technology-Based Markets

Gregory Tassey[1]

Senior Economist, National Institute of Standards and Technology

Abstract

The complexity of modern technology, especially its system character, has led to an increase the number and variety of standards that affect a single industry or market. Standards affect the R&D, production, and market penetration stages of economic activity and therefore have a significant collective effect on innovation, productivity, and market structure. Standards are classified into product-element and nonproduct categories because the two types arise from different technologies and require different formulation and implementation strategies. Because standards are a form of technical infrastructure, they have considerable public good content. Research policy must therefore include standardization in analyses of technology-based growth issues.

Keywords: standardization, innovation, R&D, economic growth, industry structure

Through R&D-performing industries and the effect of new technologies on other parts of the economy, technology accounts for one-third to more than one-half of U.S. GDP growth and at least two-thirds of productivity growth. However, the so-called "high-tech" sector only contributes approximately 7 percent of U.S. GDP.[2] This relatively small direct contribution implies substantial leverage by this sector on the overall economy, but also that extensive diffusion of new technology must take place if adequate productivity growth rates are to be achieved by the entire economy.

Standardization affects both innovation and technology diffusion. It also can influence industry structure and thereby help determine which firms benefit and which do not from technological change. Thus, a concern of R&D policy should be the evolutionary path by which a new technology

---

[1] Email: tassey@nist.gov. The author is grateful to Albert N. Link for comments on previous drafts.

[2] Tassey [1997]. The high-tech sector is defined here as consisting of four major categories: high-tech manufacturing (IT-related plus industrial electronics), communication services, software and computer-related services, and pharmaceuticals). For alternative definitions of IT-related high-tech industries, see American Electronics Association [1997, p. 128] and Department of Commerce [1998, Appendix p. A1–2]. The AEA definition results in a 6.1 GDP estimate for 1996 and the Commerce definition yields about 8 percent for 1998. To either of these definitions should be added pharmaceuticals, which brings the AEA-defined high-tech sector's GDP contribution to 7 percent.

or, more accurately, certain elements of a new technology become standardized. Over a technology's life cycle, standardization can affect economic efficiency. However, these effects can be both positive and negative. For example, standardization can increase efficiency within a technology life cycle, but it also can prolong existing life cycles to an excessive degree by inhibiting investment in the technological innovation that creates the next cycle.

Standardization can and does occur without formal promulgation as a "standard." This distinction between *de facto* and promulgated standards will be made apparent and discussed in the following sections. In one sense, standardization is a form rather than a type of infrastructure because it represents a codification of an element of an industry's technology or simply some information relevant to the conduct of economic activity. On the other hand, the selection of one of several available forms of a technology element as "the standard" has potentially important economic effects.

1. Economic Functions of Standards

A standard can be defined generally as a construct that results from reasoned, collective choice and enables agreement on solutions of recurrent problems. Looked upon in this way, a standard can be viewed as striking a balance between the requirements of users, the technological possibilities and associated costs of producers, and constraints imposed by government for the benefit of society in general (Germon [1986]).

More functionally, an *industry standard* is a set of specifications to which all elements of products, processes, formats, or procedures under its jurisdiction must conform. The process of *standardization* is the pursuit of this conformity, with the objective of increasing the efficiency of economic activity.

1.1. Nature and Scope of Impacts

Standards played an important role in the industrial revolution. They allowed factories to achieve economies of scale and enabled markets to execute transactions in an equitable and efficient manner. Standardization of parts made supplier specialization possible and increased efficiency over the entire product life cycle by facilitating part repair or replacement.

In a modern economy, standards constitute a pervasive infrastructure affecting the technology-based economy in a number of important and relatively complex ways. Some of these impacts even appear contradictory. For example, whereas the traditional economic function of standards in production can restrict product choice in exchange for the cost advantages of economies of scale, other types of standards common to advanced production and service systems can actually facilitate product variety and hence choice for the customer.

Fig. 1 depicts the multiple functions performed by standards. These functions transcend the three major stages of technology-based activity — R&D, production, and market penetration, and are difficult to construct and implement because many important technologies have both an intrinsic complexity and a "systems" character. Such characteristics demand more sophisticated technological foundations for standards and imply the need for technically competent standards setting processes.

The greater complexity of technologies and the associated networks of firms and supporting infrastructure that develop and disseminate these technologies mean that supply chains are becoming the most important level of policy analysis. Greater distribution of R&D among materials and equipment suppliers, manufacturers of products, and providers of services increasingly characterize high-tech supply chains. The consequent increase in market transactions involving technology also demands standards to reduce the associated transaction costs.

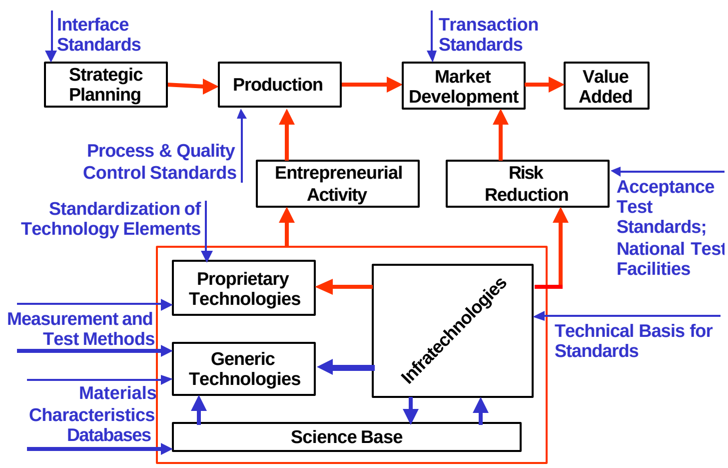

Fig. 1. Roles of Standards in a Technology-Based Industry

Technology consists of a number of discrete elements that tend to evolve in different institutional settings. These elements have distinctly different character and require different types and combinations of standards to effect efficient development and utilization (Tassey [1992, 1997]). Many variations exist but broadly defined they fall into the three major categories shown in Fig. 1:

(1) The fundamental or ***generic technology*** base of the industry on which subsequent market applications (products and services) are based

(2) A set of ***infratechnologies*** that provide a varied and critical technical infrastructure to support for development of the generic technology and subsequent market applications

(3) The market applications (***proprietary technologies***)

Because the type of R&D required for each element differs significantly, so do private-sector investment incentives with the result that underinvestment varies across the different elements of an industrial technology. The greater the infrastructural character of a technology, the more underinvestment is likely to occur. Standards and thus their technical underpinnings have a strong infrastructure character, so that underinvestment is common.

1.2. Basic Functions of Standards

To analyze the economic functions of standards in a technology-based economy, a taxonomy is required that classifies standards by functions having unique economic characteristics. For the purpose of economic impact assessment, the functions of standards are classified into four categories.[3] The positive effects of each of the four functions are described below. However, as discussed in later sections, standards also can have negative economic consequences or simply fail to achieve their maximum potential economic benefit.

1.2.1. Quality/Reliability

Standards are developed to specify acceptable product or service performance along one or more dimensions such as functional levels, performance variation, service lifetime, efficiency, safety, and environmental impact. A standard that specifies a minimum level of performance often provides the point of departure for competition in an industry. For example, a case study by Putnam, Hayes and Bartlett [1982] points out that when an automobile manufacturer develops a new engine, the company specifies the minimum acceptable lubrication attributes. This specification then becomes the basis for competition among petroleum companies, who either compete on price at the minimum specified level of quality or by offering motor oil with a level of performance above the minimum.

1.2.2. Information Standards

Standards help provide evaluated scientific and engineering information in the form of publications, electronic data bases, terminology, and test and measurement methods for describing, quantifying, and evaluating product attributes. In technologically advanced manufacturing industries, a range of measurement and test method standards provide information, which, by virtue of being universally accepted, greatly reduce transaction costs between buyer and seller. In their absence, especially for complex, technology-based products, considerable disagreement will often ensue over verification of performance claims. These disputes raise the cost of consummating a marketplace transaction, which is reflected in higher prices charged. The economic impact is to slow market penetration.

Measurement methods are also essential to conduct state-of-the-art research. In today's semiconductor R&D, scientists and engineers must be able to measure the distances between individual atoms (dopants) that are added to silicon to achieve the desired millions of high-density electronic functions on a single chip. Standardization of some of these methods is essential for the efficiency of R&D itself. For example, being able to replicate and verify research results is often critical to obtaining follow-on research funding or commitment to commercialization. Standardized scientific and engineering data (in the sense of having been critically evaluated and verified for accuracy) and standardized equipment calibration techniques are also essential for efficient R&D.

Finally, the typical manufacturing process is increasingly measurement intensive because of growing demands for quality and real-time process control. Traditional manufacturing processes tested products after a production run. The inefficiency of this approach is large, not only because of the wasted material and labor when a production run must be scrapped, but also because of down

---

[3] This taxonomy follows Tassey [1982, 1992, 1997] and Link and Tassey [1987]. David [1987] proposes a similar taxonomy based on three kinds of standards (reference, minimum quality, and compatibility). Other taxonomies have been developed based on the process by which a standard comes into existence. For example, David and Greenstein [1990] provide a framework to classify standards as *de facto* ("unsponsored" or "sponsored") or promulgated (voluntary or *dejure*).

time and hence unused capacity incurred while a problem is identified and fixed.[4] The availability of computers makes possible the real-time monitoring and control of a process, potentially enabling instant adjustment of process variables.

When fully implemented, real-time control can virtually eliminate waste and increases product mix flexibility. However, real-time control of a production process is a complex systems technology, requiring large numbers of sensors, computers, and software. Standardization of certain performance measurements for, say, a sensor facilitates design of the equipment. Equally large efficiency gains can occur from standardization of certain elements of the actual methods or techniques of process control, as adherence to these standards often removes the need for much of the post production testing. Finally, whatever production strategy is adopted, the equipment that helps execute the strategy must be periodically calibrated by standards in order to ensure maximum efficient performance.

1.2.3. Compatibility/Interoperability

Standards specify properties that a product must have in order to work (physically or functionally) with complementary products within a product or service system. This function of standards has been the most intensively studied by economists.[5] Compatibility or interoperability is typically manifested in the form of a standardized interface between components of a larger system. An effective interface standard does not affect the design of the components themselves, such as numerically controlled machine tools or the components of these tools, including controllers. In fact, interface standards provide "open" systems and thereby allow multiple proprietary component designs to coexist — that is, they enable innovation at the component level by being competitively neutral with respect to design. In effect, competitors can innovate on "either side" of the interface, while the consumer of the product system can select the particular components that optimize system design. They also allow substitution of more advanced components as they become available over time, thereby greatly reducing the risk of obsolescence of the entire system. Widespread factory automation as it is currently evolving in advanced economies likely will not proceed without these standards.

Without interface standards, large companies often supply "turnkey" systems where proprietary interfaces link components. However, the cost to the user can be high because the system is not optimized for the user's particular needs (competitors will typically offer components that are superior to some of those in the turnkey system) and price competition will not be a factor when system components need replacement.

In such situations, system design can still be optimized. However, the cost of modifying physical and functional interfaces to allow components from different vendors to work together (i.e., to "interoperate") is usually prohibitive. Moreover, full functionality is often not obtained by reengineering proprietary (nonstandard) interfaces.

---

[4] Scrap can result when periodic testing of product reveals an attribute flaw. All units produced between the sample product tested and the previous test are equally defective.

[5] See, for example, David and Steinmueller [1994], Kahin and Abbate [1995], and Link and Scott [1998].

1.2.4. Variety Reduction

Standards limit a product to a certain range or number of characteristics such as size or quality levels. The fourth function of standards is the traditional one of reducing variety to attain economies of scale. The majority of standards perform this function. However, variety reduction is no longer simply a matter of selecting certain physical dimensions of a product for standardization (such as the width between threads of a screw). Variety reduction is now commonly applied to nonphysical attributes such as data formats and combined physical and functional attributes such as computer architectures and peripheral interfaces.

The process of setting variety reduction standards also varies significantly. Many standards of this type are viewed as infrastructure and thus adopted by an industry consensus process. However, standardization of some attribute or element of a product is just as often achieved through the marketplace by one firm gaining control of the underlying technology and using this control to force other manufacturers with whom that firm competes to adopt its version of the technology. The product element then becomes a de facto (non-consensus) standard.[6]

Conceptually, the variety reduction function is the most difficult category of standardization to analyze because of its ability to either enhance or inhibit innovation. Variety reduction typically enables economies of scale to be achieved, but larger production volumes tend to promote more capital-intensive process technologies. This common evolutionary pattern of a technology over a number of product life cycles usually reduces the number of suppliers and increases their average size. Such trends may or may not reduce competition, but often progressively exclude small, potentially innovative firms from entry due to increased minimum efficient scale thresholds.

2. Types of Standards

Standards have been classified by the form they take. To varying degrees all four functions of standards specified above can be described alternatively in terms of design specifications or performance levels. Design-based standards are much more restrictive and can inhibit innovation to a greater degree than performance-based standards. The latter type allows flexibility in product or service design while still meeting the performance requirements of the standard. Thus, standards generally work more efficiently when they are performance based.

The economic functions of standards aggregate into two basic categories — product and non-product — delineated by their relationship to product (or service) structure and their public good content. Distinctions between the two types are important because their economic roles are distinctly different and hence so are the processes by which they are set. The differences in public good content have important implications for policy because the rationales for government intervention are very different in the two cases.

Simply put, standardization of one or more attributes of a product (or service) can convey direct competitive advantage to the owner/controller of the technology producing those attributes. They

[6] Intel's microprocessor architecture and Microsoft's operating system are well-known examples.

therefore get considerable attention from economic growth strategists. Conversely, non-product standards tend to be competitively neutral, at least within an industry or trading block. Hence, they tend to get less attention. Yet, this latter category can be critical to the entire industry's efficiency and its overall market penetration rate.

2.1. Product-Element Standards

Product-element standards typically involve one of the key attributes or elements of a product, as opposed to the entire product. Government, especially where large economies of scale are present or when early market entry is considered an essential part of a national economic strategy, mandates some product-element standards.

In most cases, at least in the U.S. economy, market dynamics determine a *de facto* standard. Alternative technologies intensely compete until the dominant version gains sufficient market share to become the single standard. Market control by one firm can truncate this competitive process. Such control is particularly effective in cases of increasing returns and can quickly force acceptance of the monopolist's proprietary technology element as the standard. However, the globalization of high-tech markets with shorter product life cycles is making single-firm dominance more difficult. In response, various combinations of vertical and horizontal consortia are promulgating product-element standards by consensus, at least within a single large economy or trading block.

One of the more visible examples of the competitive effects of *de facto* standardization is the "architecture" of personal computers.[7] Elements of computer architecture such as the operating system, the "bus," the graphical user interface, and the applications programming interface, have been the focus of intense competition by firms seeking to gain sufficient market control to "set" the *de facto* standard for the particular product element.[8]

In this regard, Apple Computer made a brilliant move when it forced third-party software developers to adopt a standard "graphical user interface" so that all programs running on Apple's computers presented the user with the same screen format and command structure.[9] However, Apple kept its hardware operating system proprietary (i.e., it did not adopt an open systems architecture strategy) and thus it had no chance to become the industry standard. This decision explains Apple's initial prosperity and subsequent competitive decline. In contrast, Sun Microsystems opened the microprocessor architecture for its workstations in order to obtain help in gaining market share. It

---

[7] Architecture is the scheme by which the function of a product is allocated to physical components (Ulrich, [1995, p. 419]). Architecture is important because it provides a set of standardized product attributes and the rules or protocols for their interaction with other product elements.

[8] The business literature continues to debate the efficacy of competition to supply the product elements that constitute an architecture. See Morris and Ferguson [1993].

[9] The term *graphical user interface* refers to the use of "menus" of images instead of characters to indicate instruction options for the computer user. Apple pioneered this much-preferred format and also made this approach even more attractive to users by sticking to a standardized format for all application programs.

was willing to share its markets in return for the increased probability that its architecture would become a standard and thereby create a large and stable customer base.

The policy issue is thus: To the degree that certain product elements must become standardized to enable economies of scale and network externalities to be realized, the potentially large benefits that accrue to the firm owning the technology element are acceptable. However, this is the case as long as competition is not constrained in the other elements that make up the overall product or system technology.

This caveat with respect to economic efficiency was the basis for charges by software firms that Microsoft was using its monopoly position as owner of the dominant operating system standard to gain unfair advantage in marketing its own versions of new operating system elements (such as a web browser) and applications programs that run on that operating system. Such charges were pursued in a 1998–1999 U.S. government antitrust lawsuit.

In contrast, the Japanese PC industry has never been able to establish a single standard operating system. Reasons for this include (1) an oligopolistic industry structure (*keiretsu*), which has resulted in factionalism within the industry and thereby prevented a single standard (domestic or foreign) from gaining dominance; (2) a mainframe computer orientation, which for a long time relegated PCs to basically terminal status, thereby slowing market growth; (3) a reluctance to move away from customized software; and (4) a language barrier—the inability to export kanji-based software (Cotrell [1994]).

Over a technology's life cycle, additional elements of the product's technology become standardized so that the product takes on a "commodity" character. Competition among suppliers of the "standardized" product then becomes increasingly based on price and service-related aspects of the product's acquisition and use. Dell Computer, for example, has succeeded in the PC industry by being the low-cost producer and offering excellent before and after sales service. This evolutionary pattern was noted decades ago by the famous Austrian economist Joseph Schumpeter [1950], who observed that one of the essential dynamics of capitalism is assuring that the "silk stockings" initially purchased only by the rich would eventually be items of mass consumption.

2.2. Nonproduct Standards

Nonproduct standards, as the name implies, derive from a different technical base than that upon which the attributes of the product itself depend. Industry organizations often set these standards using consensus processes. The technical bases (infratechnologies) for such standards have large, although not total, public good content, so that their provision frequently depends upon a combination of industry and government investment. Examples of infratechnologies frequently embodied in nonproduct standards include measurement and test methods, interface standards, scientific and engineering databases, and artifacts such as standard reference materials (Tassey [1997, Chaps. 8 and 9]).

Most infratechnologies and therefore the resulting industry standards are derived from basic standards. Basic standards represent the most accurate statements of the fundamental laws of physics and have such diverse applications that they qualify as pure public goods and hence are provided entirely by government. Basic standards are relatively few in number and are not easily transported to or used by industry. Thus, as indicated in Fig. 2, such standards are converted into working and transfer standards that convey the standardized information to industry. Large numbers of industry standards are based on (traced back to) basic standards. With respect to form, most industry standards are either method, procedural, or normative.[10]

Fig. 2 provides an example of the hierarchy through which basic standards are utilized to develop infratechnologies upon which semiconductor industry standards are based. The production of semiconductor components is a highly demanding process. The densities of today's circuits are such that each conducting path ("line") on a chip is a small fraction of the width of a human hair. These widths must be consistent with respect to design specifications to avoid thermal, electrical, and other problems. The semiconductor producer therefore needs to be able to measure the widths of circuit lines that make up a "chip." Particularly important are line widths on the masks that are used to inscribe the multiple layers of circuit patterns on the chip itself. Such masks are used to make tens of thousands of chips. Their quality greatly affects performance of the chips produced and hence a semiconductor manufacturer's production yield.

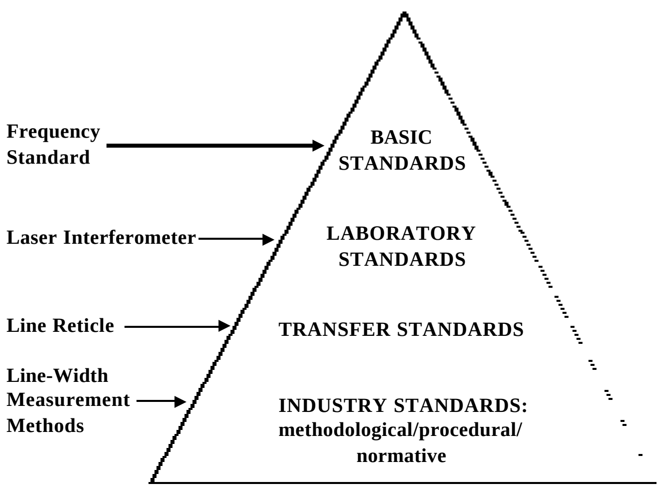

Fig. 2. Hierarchy of Nonproduct Standards (Tassey, Chap. 9)

The line-width measurement equipment must be calibrated against a physical standard, which has a pattern of lines whose thickness and spacing have been determined to a specified level of accuracy. This determination is done by an authoritative source, such as the National Institute of Standards and Technology (NIST) in the United States. The physical or "transfer" standard used by industry must be easily transportable (a reticule in the above example) in order to ensure widespread and accurate transfer of the infratechnology.

The information transferred by a physical standard is itself determined or certified by a so-called working standard, which is laboratory based and more accurate but not readily transportable. In this example, the working standard is a laser interferometer, which measures and certifies the physical

[10] A normative standard is one in which a particular value (size, performance, quality, or design) is selected from a range of alternative values.

dimensions of the line reticule prior to transfer to industry. Finally, the laser interferometer is itself dependent for calibration on a "basic" standard for length. [11]

The four basic functions of standards are all represented in the range of standards for a single industry like semiconductors. Many standards, like line-width measurement, provide information, while others affect variety reduction, quality, reliability, or compatibility. The collective economic impact of these standards is greatly magnified when an entire supply chain is considered. Semiconductors are a component of computers and communications equipment, which, in turn, constitute an information network. Each level in the supply chain requires an elaborate infrastructure of standards.[12]

3. Market Structure Effects of Standards

The market structure effects of standardization have important effects on the achievement of economic growth objectives. “Open systems” allow small and medium companies to participate in markets for system technologies by supplying components in which they have a comparative advantage. This diversification on the supply side of the market makes system optimization by users (the demand side) possible and increases price competition.

Decisions by large companies to enter systems markets at the component level are also affected by the availability of interface standards. These firms typically target larger markets where they can benefit from economies of scale. A strategy of selling into markets characterized by proprietary turnkey systems requires a high degree of product segmentation in order to service each turnkey system integrator. This situation increases costs and constrains market growth projections, even with a superior technology. The more distributed the participants in a market, the more critical to technological innovation are open systems.

A good example is medical services, which has traditionally suffered from excessive customization. Moreover, healthcare environments are becoming extremely distributed. In every

---

[11] Line-width measurement is just one of many infratechnologies that a competitive semiconductor industry must utilize. For example, current state-of-the-art chips consist of multiple layers of circuits. The circuits in each layer must be connected to adjoining layers. Accomplishing this very difficult manufacturing step requires a precise alignment of the “mask” (circuit pattern) for each layer. Until recently, the alignment process required multimillion-dollar optical equipment. NIST, however, developed a procedure that allows semiconductor manufacturers to ensure proper alignment of successive layers of an integrated circuit with a precision better than ten nanometers. This new calibration standard represents a more than fivefold improvement over current alignment calibration methods and is much less expensive. The cumulative economic impact of such advances in infratechnologies/standards is substantial and can greatly affect price as well as quality, and hence competitive position for the domestic industry.

[12] In communication networks, for example, there are published standards and interface protocols that allow hardware components and software from many vendors to operate as a single product or as a system of products (i.e., as a network). To achieve efficient or "seamless" integration, the standards and protocols define what rules hardware components must adhere to in order to exchange signals between applications software and operating systems at different levels in the network.

local community there are multiple sites — hospitals, labs, clinics, pharmacies, home-care settings, and so on — that conduct numerous clinical and administrative activities. Currently, the operational computing infrastructures across these organizations are rarely the same. In this environment, incomplete standardization is an increasingly serious problem. User-defined attributes for specific applications make plug-and-play data interchange impossible. Even where two applications are built according to an existing standard, frequent user-defined modifications to the applications often prohibit "seamless" communication between them.

The solution to such a problem, where standards exist but customization of applications creates incompatibility with the existing standardized interface, is to develop and reach agreement on true open systems. Achieving complete interoperability eliminates levels of complexity in implementing limited or partial standards. In the healthcare area, Hewlett-Packard took the lead in organizing an industry consortium (The Andover Working Group) to build an "enterprise communication framework (ECF)" that provides a common interface for data communications. HP was motivated by its desire to sell multiple products into a single large market. With interoperability among system components, such a market retains the advantage of diversification at the component and system levels and also achieves the efficiency advantages of interoperability.

Such interoperability is a demanding challenge. It is achieved by encapsulating existing standard message profile specifications into software objects called enterprise communicators. These communicators provide a comprehensive set of rules that make it possible for developers to achieve interoperability for their applications without having to completely understand the existing complex standards and rewrite them into their applications. This software code is packaged within the ECF and embedded into third-party applications as a software component.

In practice, attaining truly open computer systems has been extremely difficult, with achievement of compatibility being partial at best. One of the major reasons for this difficulty is the complexity of systems technology. Each one of a number of layers in the hierarchical structure of a typical information system must have its own architecture and these architectures must be integrated into a larger network. Today, systems technologies not only have complex hardware and software platforms, but require rules and protocols to introduce, process, store, move, and retrieve information.

The existing systems architectures are typically controlled by a handful of firms. Vendors often claim to conform to open systems' requirements, but in reality they conform only to a portion of the set of standards to which compliance is necessary to sell their systems. Moreover, even when true open systems compatibility is offered, users often find that mobility from their current proprietary system to the open system does not exist, even when the same vendor provides both. That is, mobility products are not made available, leaving the user with a severe transition problem. Such partial conformity standards tend to segment the market structure.

In summary, lack of conformity to interface standards by the supply side of a market can be a conscious strategy or a reflection of the poor quality of the standard.[13] Large firms with market strategies focusing on turnkey systems have an incentive to resist open systems to protect the market

[13] See David and Greenstein [1990] for an excellent review of the early literature on these topics.

shares resulting from their horizontally integrated strategy. In cases of increasing returns, such strategies may result in a monopoly position being attained by one supplier or prolonged market segmentation by several competing suppliers. Either situation has the potential to constrain economic efficiency.

Conversely, large users are likely to want to have access to several different classes or brands of a particular product, or at a minimum have a second source of supply for a critical component. Such product users have substantial need for effective interfaces. The ability to integrate equipment and software from different vendors (or even to integrate different classes of equipment from a single multi-product vendor) has several economic benefits: system costs can be reduced and performance optimized thereby increasing productivity, competition among vendors for each element of the system lowers prices, and responding to technological or functional obsolescence is facilitated by replacing individual components at relatively low cost.

At the product or component level, increasing returns can result in a monopoly position for one supplier. If the product happens to be critical in some way, that monopoly can be extended into related component markets. In other cases, multiple standards exist for extended periods of time resulting in a segmented market structure. The latter situation is caused by high switching costs relative to perceived benefit gains by users. These switching costs can be increased when suppliers with monopolistic control of portions of the market for the main product (such as a minicomputer) fail to agree on a single standard for a complementary product (such as the operating system). The UNIX operating system is an important example.

4. "Systems" Aspects of Standards

Product and nonproduct standards within a product or service system interact with each other to create a systems character to the overall economic effects of standardization. The interactions among different types of standards that affect a single industry are summarized in Fig. 3. Each of the two industries shown here depends on standardized formats and evaluations of scientific and technical information from supporting science and technology infrastructures. They use this information to select and conduct R&D projects.

These industries then set additional standards specific to their internal needs for production efficiency. Finally, each industry must also be concerned with standards that affect their ability to sell their products in specific markets. Meeting procurement specifications in domestic markets or complying with certification requirements for exporting to foreign markets are examples. Moreover, as many products now become components of larger product systems, part of meeting specifications is demonstrating compatibility with the performance attributes of other products.

Interaction among interface standards occurs because of the need to optimize product systems. For example, a numerically controlled machine tool is a system of components such as the controller and numerous sensors. Link and Tassey [1987, pp. 88–96] point out that the multiple interface standards involved in a machine tool allow modularization of the product, which permits custom design and prevents obsolescence. However, these standards must be linked by a common data format, if all components are to function together as an efficient system.

Moreover, the structure of the components affects the interface standards. The technical difficulty and hence cost of constructing an interface standard increases when all components are completely proprietary. A frequent solution is to standardize certain attributes or elements of key components. In the case of the machine tool, a standard architecture for the machine tool controller has been agreed upon. This step has significant efficiency advantages, such as allowing upgrades to the controller without costly reengineering of the interfaces between the controller and other hardware and software components.

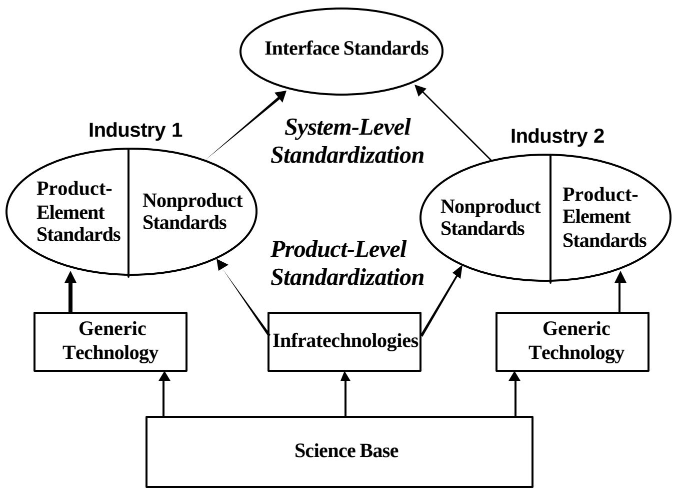

Fig. 3. The Structure of Industrial Standardization [Tassey, Chap. 9]).

As Tassey [1997, pp. 172–175] points out, industries that have complex, measurement intensive production processes must reengineer multiple standards simultaneously for each new generation of the generic product technology. The semiconductor industry is an excellent example. In the past, a single dominant integrated circuit supplier has undertaken the development of standards associated with transitions between generations of semiconductor process technology (usually revolving around wafer size).

However, accomplishing the transition now requires a complex and integrated standards development process, involving a large number of different standards. The huge estimated cost of next generation fabrication plants and shorter product life cycles finally have made the risk too great for individual IC manufacturers to incur the considerable burden of undertaking the standards development efforts alone. As a result, an international consortium of semiconductor firms is undertaking the development of standards for the generation of semiconductors based on a 300 millimeter wafer.

5. Potential Negative Impacts of Standardization

Both types of standardization described here can greatly increase overall economic efficiency and hence rates of growth in technology-based industries. However, through their structure or through the timing of their implementation, they can also have negative economic impacts.

5.1. Product-Element Standards

Once a product-element standard is set, network externalities are typically realized. However, the marketplace dynamics that result in one firm's version of the technology becoming the standard do not guarantee that this version is the optimal one. A "lock-in" effect ensues, whereby many developers of related products conform to the standard and purchasers of the standardized technology invest substantial resources in learning to absorb and use it as well as complementary technologies and infrastructures. These sunk costs create a reluctance to switch to a new standard and related cluster of technologies. Considerable resources are then allocated to an inferior technology, which can extend over long periods of time.[14]

As pointed out by a number of economists, the controller of the product-element standard captures increasing returns to scale as initial market penetration begets dominance. This may be an acceptable price for rapid market growth from an economic welfare perspective, as long as either the standardized version of the technology element is the best one or sufficient opportunity exists for market forces to periodically replace the standard.[15]

The tendency for modern technologies to have systems structures increases the economic importance of the lock-in phenomenon. In particular, most technologies today have clusters of standards embedded in them. One, however, is frequently the dominant or driver standard. Thus, the controller of that standard can have even greater impact on the evolutionary development of the overall technology and related markets.

While the lock-in effect works on the demand side, the "installed base" effect impacts the supply side. Dominant suppliers of a technology, who either control or adhere to the standard, have invested substantial resources in developing and servicing the markets based on this standard. Their established market positions promote evolutionary as opposed to revolutionary migration of customers, who demand backward mobility to existing technology. Mobility is more difficult and expensive, the more radical the new technology. If Microsoft, Apple, and IBM were freed from maintaining compatibility with existing hardware software, new generations of systems software would likely evolve at a faster pace.

5.2. Nonproduct Standards

Because this class of standards consists of a set of "tools" that make the process of developing, producing, and marketing the core technology more efficient, they do not have the dominant or controlling effects of product-element standards. Thus, the potential negative impacts are not as obvious. However, negative economic effects can occur. They typically result from three problems: (1) multiple standards for the same infratechnology, (2) poorly designed standards, and (3) poorly timed standards.

Multiple standards usually arise in the early phases of a technology's development when innovating firms are intent on achieving market penetration. At this point, the market is often not

[14] For more than a decade, DOS — a clearly inferior operating system to alternatives such as Apple Computer's system — dominated the personal computing market because the Microsoft/IBM combination was able to lock in the PC market.

[15] Arthur [1996] provides an excellent analysis of the increasing returns-to-scale phenomenon that characterizes many technology-based industries and the consequent motivation for firms to set and control product-element standards.

large enough to warrant investment in infratechnologies such as sophisticated test methods. Consequently, innovators rely on off-the-shelf or contrived methods that hopefully convince customers to try the new technology.

However, as the market expands, imitators enter with their own versions of the required test methods. Confusion among buyers soon takes over, as different vendors provide test methods that produce different results. The cost to buyers of determining what test method most accurately measures performance rises as their purchases increase. Moreover, because these multiple measurement and test methods are typically adapted from previous generations of the underlying infratechnology, they are frequently inaccurate and sometimes markedly so.

These first two problems can be resolved by additional research to develop better infratechnologies, which can then be adopted as single industry standards. However, because a consensus process involving numerous industry stakeholders promulgates most standards of this type, developing the needed infratechnology and reaching a consensus on the standard often encounters substantial obstacles.[16] The resulting delays can create significant costs in executing market transactions, which raise prices and thereby retard market penetration.

The problems of multiple standards and the timing of standardization also arise at the interface between components of a system of technologies. The increasingly important objectives of compatibility and interoperability are implemented through interface standards. Interface standards typically have both physical and functional elements and can be extremely complicated to design, especially in systems technologies where multiple components must interoperate.

The timing of interface standards is not only important, but also often essential for small firms that supply a single component to a larger system. Often, a standard interface must be in place for small firms to even contemplate entering the market. Thus, poorly designed or delayed interface standards can have decidedly negative effects on market structure.[17]

5.3. Cost

The economic cost of inadequate standardization (either no standard or a poorly designed one) can be very high. A study by Brunnermeier and Martin [1999] estimated that interoperability problems associated with sharing product and engineering data impose annual costs on the U.S. automotive supply chain totaling approximately $1 billion. Resources expended to correct or re-create data files because of software incompatibilities account for more than 80 percent of the total. Delays in the introduction of new vehicles are responsible for almost 10 percent of the cost. Other expenses include purchases of several vendors' versions of software designed to perform similar tasks and spending for data-exchange services. Other large U.S. industries, such as aerospace, have similarly large electronic data transmission requirements within their supply chains, so the total cost

[16] See U.S. Congress, Office of Technology Assessment [1992].

[17] Delays can also occur simply because all component technologies do not evolve at the same rate. Laggard components of the system are often characterized by chaotic entry and exit as well as by rapid technological change, which prohibits consensus being reached on needed standards.

to the U.S. economy of inadequate interoperability standards is much greater than the estimate for the automotive supply chain alone.

One potential solution is the evolving international Standard for the Exchange of Product Model Data (STEP), or ISO 1303. STEP is a neutral file format intended to support computer-to-computer exchanges of all types of product data, from initial design to maintenance requirements. However, this standard is itself very complex and is requiring a huge investment worldwide to develop the generic set of standards. In addition, infratechnologies such as test methods and software tools must be developed to facilitate industry's adoption and implementation of interoperability standards.

6. Degrees of Standardization

From both the corporate strategy and public policy points of view, standardization is not an all-or-nothing proposition. In complicated system technologies, such as distributed data processing, telecommunications, or factory automation, standardization typically proceeds in an evolutionary manner in lock step with the evolution of both embodied and disembodied technology. The pattern of evolution is determined by several factors: the pace of technological change embodied in each component category; disembodied technology development, which determines the overall system architecture and organization; and changes in market structure (and, with it, the incentives and ability to force the standardization process).

In the case of numerically controlled machine tools, early total standardization of data formats would have severely compromised the range of performance attributes desired by different users in the machine tools they purchased. Thus, a degree of standardization has been optimal, at least up to this point in the technology's evolution. In other words, complete standardization too early in the technology life cycle can constrain innovation.

As technology-based systems become increasingly important and "windows of opportunity" for making successful investments in the associated markets continue to shrink, the relevant standards will have to be managed. If a standard is fixed, even if it is competitively neutral, it will eventually act to stifle the introduction of new technology into the system. Alternatively, if the standard is updated frequently, then version consistency (upward mobility of current system components) can become a problem.

For example, efficient data processing and communications networks are possible only if standard interfaces are provided on all the communication paths in the network. Such interfaces need to be defined between application programs, data formats, network protocols, printer control codes, human/machine interfaces, and so on. But, these system elements are all evolving at different rates and thus need updated interfaces at different points in time.

Thus, with new technologies continuously being introduced into ever expanding networks, the pressure on the standards infrastructure to adapt is substantial. One reason is the need to reflect changes in the core distributed processing technology. Another is the requirement to adapt system management technologies such as toleration of faults, rapid (real time) response to multimedia inputs and outputs (voice, image, text, and others), and human friendly interfaces to a constantly changing network structure as nodes are added, removed, or replaced by different systems.

7. Policy Implications

In most cases, market dynamics seem to affect standardization of product elements that drive network externalities. As long as the private-sector originator and hence controller of the standard does not use this control to gain advantage in related products, long-run efficiency seems attainable. On the other hand, if monopoly control by a single firm becomes entrenched or, conversely, if no standard emerges (i.e., multiple standards persist) to the detriment of economic growth, some policy action may be required.

Cottrell [1994, pp. 163–164] discusses how the Japanese government tried in several ways to remedy the multiple PC standards problem in its domestic industry. The SIGMA (Software Industrial Generalization and Maintenance Aids) Project and a subsequent five-year extension (SuperSIGMA) were major government efforts to achieve a Unix operating system standard and thereby deal with the economic inefficiencies of a heterogeneous (nonstandardized) product structure. The project attempted to promote automated software development and create standardized Unix-based data structures, communications interfaces, and applications platforms.

An important policy implication is that research projects can be funded with a pure "technology development" objective or a combined "technology and standardization" objective. The TRON project in Japan is an example of the latter.[18] R&D projects that involve horizontal cooperation between suppliers of system components (e.g., computer hardware and software firms) can promote standardization of key product elements at an earlier point in the technology life cycle. Such joint research ventures with standardization as an objective are occurring more frequently, especially in IT industries.

Finally, government R&D can establish and demonstrate a backbone infrastructure, which in turn promotes private-sector R&D investment in standards to enable effective use of this infrastructure. No private entity would have funded several decades of the Internet's development, as it slowly evolved to its current state where enormous private investment is now taking place. Industry standards such as Java are enabling a host of applications to be developed and implemented, but Java would not have been possible without the U.S. government's persistent funding of the Internet's predecessors.

This last example exemplifies a key issue for government R&D policy. Success as an innovator manifests itself in a large installed base. The lock-in effect that results from the role of increasing

[18] The TRON Project has been an ongoing effort between Japanese industry and universities since 1981. It seeks to develop standardized operating systems, which in turn will drive compatible hardware development. However, the installed base effect of the dominant U.S. industry standards has at least partially thwarted this effort, as one of the largest Japanese computer firms, NEC, has invested so much effort in marketing MS-DOS/ WINDOWS/UNIX machines that the company has not participated in TRON. See Cottrell [1994, p. 166] and Kahaner [1991].

returns to scale can be accepted as a natural and even desirable attribute of many technology-based markets. However, this characteristic along with the installed-base effect can combine to prolong technology life cycles at the expense of newer, more efficient technologies. The standard setter can maintain control by migrating the installed base through several incremental generations of the technology within a longer life cycle. This supply-side incentive to resist major technological change is reinforced by the fact that, once a large customer base is attained, user demand for backward mobility further promotes an evolutionary rather than a revolutionary approach to change. This phenomenon is a major reason why smaller firms appear more innovative, especially with respect to radical innovations; they are not encumbered by the installed base.

One possible policy response is to require, through regulation, equal access to applications markets so that the firm controlling the infrastructure standard does not use this advantage to restrain competition (Arthur [1996, p. 106]). Government funding of new generations of technology-based infrastructure is another response to dealing with the negative effects of this type of standardization.

The Internet provides a great example of the basis for a new technology life cycle that may eventually replace the "Wintel" standards of the current PC technology life cycle.[19] In fact, the "open systems" nature of the Internet may preclude the type of lock-in strategy observed so often in the 1980s. Netscape gave away its Internet browser and attained 70 percent of the market. Its objective was to obtain potentially larger and more lasting profits from the substantial value added available from software applications accessed through the browser. Similarly, America Online gained an initial substantial lead of more than 4.5 million subscribers by giving away free services. In these and other similar cases, however, ease of entry implies achieving lock-in is problematic at best.

Policy implications for non-product standards and investment in the associated infratechnologies are subtler. Public sector infratechnology investment is not resource intensive. The semiconductor industry, for example, spends $4-5 billion per year on measurement equipment (Finan [1997]). However, the government support for the infratechnologies (measurement and test methods, for example) and associated standards required to attain the productivity goals of this investment is considerably less. Nevertheless, tight R&D budgets in most industrialized nations during the 1990s have constrained infratechnology research support.[20]

The critical point is that a range of infratechnologies and associated standards are needed by industry over the technology life cycle. Individual standards may have modest economic impact, but the large number of standards required by the typical high-tech industry means that the collective impact is substantial. Government research program management and associated funding must be sufficient to provide this complex technical infrastructure in the time frames dictated by industry R&D strategies in response to competitive trends.

---

[19] "Wintel" refers to the dual hardware and software standards controlled by Intel and Microsoft, respectively.

[20] For example, NIST infratechnology research funding relative to U.S. industry R&D spending, which NIST is charged with supporting, is one half the level of 25 years ago.

8. Conclusions

Over the past decade, the infrastructure roles of standards have increased in importance because (1) many new technologies are systems or networks so that increasing returns to scale can generate huge economic rewards for the version of a technology that becomes the standard, (2) the demand for quality and reliability in technologically complex products and systems requires a range of standards based on sophisticated infratechnologies, (3) the systems nature of critically important technologies means that competition is greatly affected by the degree of standardization within product structures and at the interfaces between components of these systems, and (4) the shortening of the average technology life cycle has on average increased the pressure on the standards setting process with respect to timing.

The central strategic problem of managing the timing and content of standards is a difficult one because (1) many types of standards are needed in today's typical technology-based industry, (2) they interact to varying degrees with one another, and (3) nonproduct standards, as one type of technology infrastructure, derive from different sources than the industry's core technology and thereby often conflict with corporate strategies.

Because the U.S. domestic market is large with considerable internal competition in most industries, American economic philosophy has allowed the competitive dynamics of the marketplace to set product-element standards. Issues of unfair advantage in related markets for the firm whose technology eventually wins out as the industry standard are raised periodically. In other cases, several competing “local” standards coexist for some time, resulting in complaints of inefficiency. In Europe, more coordinated standardization efforts in areas such as wireless communications have been undertaken to gain first mover advantages. This latter approach is causing U.S. industry and government managers to rethink their laissez faire standards strategies.

More generally, industry structures and hence long-term economic efficiency can be significantly affected by which firms’ technologies become industry standards and the degree to which standardization enables or inhibits access to the markets for a technology. To the extent that monopolistic control of a standard or the existence of multiple standards produces economic inefficiency, the globalization of technology-based competition can accentuate these problems.

More national economies are now capable of competing for the dominant product-element standard in technology-based industries. National governments often support the domestic industry’s standard over alternatives. Moreover, multinational alliances of firms frequently pursue their versions of a product-element standard in emerging technology-based markets. Thus, multiple standards can arise and persist for some time. In such cases, multilateral efforts should be undertaken to at least harmonize these competing standards, if not to select one as a single international standard at an appropriate point in the technology life cycle.

In contrast, nonproduct standards have a more persuasive infrastructure character and thus, unlike product-element standards, evolve largely through non-traditional market processes. Such standards are themselves based on infratechnologies, which often derive from a different science and engineering base than does the industry’s core technology. If correctly configured, they are competitively neutral and have their economic impact by increasing efficiency for all domestic suppliers and customers in the particular market for the product or service.

The unavailability of standards at different points in a technology's life cycle can result in large economic inefficiency. Both types of standards can also cause economic losses if they are poorly structured. Multiple standards may exist for prolonged periods of time, limiting economies of scale or network externalities and hence total market growth, although market growth may be greater in the long run if this situation allows superior technology to eventually dominate. Therefore, especially because the process of standardization within a technology life cycle is often irreversible, these and other elements of the dynamics of this process are key policy variables.

Finally, standardization over a technology life cycle has a dynamic character, in that different degrees of standardization are optimal at different points in both the technology's and the industry's evolution. Moreover, because standards interact with each other in imparting their economic impacts, the process of standardization frequently must be managed as a system.

References

American Electronics Association [1997], *Cybernation*. Washington, DC: The American Electronics Association.

Arthur, W. Brian [1996], "Increasing Returns and the New World of Business," *Harvard Business Review* 74 (July–August): 100–109.

Brunnermeier, Smita and Sheila Martin, *Interoperability Cost Analysis of the U.S. Automobile Supply Chain* (NIST Planning Report 99–1). Gaithersburg, MD: National Institute of Standards and Technology.

Cotrell, Tom [1994], "Fragmented Standards and the Development of Japan's Microcomputer Software Industry," *Research Policy* 23: 143–174.

David, Paul A. [1987], "Some New Standards for the Economics of Standardization in the Information Age" in P. Dasgupta and P. L. Stoneman (eds.), *The Economic Theory of Technology Policy*. London: Cambridge University Press.

David, Paul A. and Shane Greenstein [1990], "The Economics of Compatibility Standards: An Introduction to Recent Research," *Economics of Innovation and New Technology 1*: 3–41.

David, Paul A. and W. Edward Steinmueller [1994], "Economics of Compatibility Standards and Competition in Telecommunications Networks," Information Economics and Policy 6 (December): 217–241.

Department of Commerce [1998], *The Emerging Digital Economy*. Washington, DC: U.S. Department of Commerce (April).

Finan, William [1997], *Metrology-Related Costs in the U.S. Semiconductor Industry, 1990, 1996, and 2001* (Planning Report 98-4). Washington, DC: Technecon Analytic Research (prepared for the Program Office, National Institute of Standards and Technology (May).

Germon, C. [1986], "La normalisation, cle d'un nouvel essor," la documentation francais (report to the Organization for Economic Cooperation and Development). Paris: OECD.

Horwarth, R. and G. Lee [1996], "300 mm: Not 'If' but 'When' and 'How'," *Channel* (June–July).

Japanese Industrial Standards Committee [1990], *Recommendations for a Long-Range Plan for the Promotion of Industrial Standardization* (June 5).

Kahaner, D. [1991], "TRON (The Real Time Operating System Nucleus)," *Scientific Information Bulletin* 16 (3): 11–19.

Kahin, Brian and Janet Abbate [1995] (eds.), Standards Policy for Information Infrastructure. Cambridge, MA: The MIT Press.

Link, Albert N. and Gregory Tassey [1987], *Strategies for Technology-Based Competition*. Lexington, MA: Lexington Books.

Link, Albert. N. and John. T. Scott [1998], "Assessing the Infrastructural Needs of a Technology-Based Services Sector: A New Approach to Technology Policy Planning," *STI Review*, No. 22: 171–207.

Morris, C.R. and C.H. Ferguson [1993], "How Architecture Wins the Technology Wars," *Harvard Business Review* 71 (March–April): 86–96.

Office of Industries, U.S. International Trade Commission [1998], *Global Assessment of Standards Barriers to Trade in the Information Technology Industry* (Publication 3141). Washington, DC: U.S. International Trade Commission (November).

Putnam, Hayes and Bartlett [1982], *The Impact of Private Voluntary Standards on Industrial Innovation*, vol. 2. Gaithersburg, MD: The National Bureau of Standards.

Schumpeter, Joseph [1950], *Capitalism, Socialism, and Democracy*. New York: Harper and Row.

Tassey, Gregory [1982], "The Role of Government in Supporting Measurement Standards for High-Technology Industries," *Research Policy 11*: 311-320.

Tassey, Gregory [1992], *Technology Infrastructure and Competition Position*. Norwell, MA: Kluwer.

Tassey, Gregory [1997], *The Economics of R&D Policy*. Westport, CT: Quorum Books.

Tassey, Gregory [1999], *R&D Trends in the U.S. Economy: Strategies and Policy Implications* (Planning Report 99–2). Gaithersburg, MD: The National Institute of Standards and Technology.

Ulrich, K. [1995], "The Role of Product Architecture in the Manufacturing Firm," *Research Policy* 24 (May): 419–440.

U.S. Congress, Office of Technology Assessment [1992], *Global Standards: Building Blocks for the Future* (TCT-512). Washington, DC: U.S. Government Printing Office.