THE ECONOMICS OF COMPATIBILITY STANDARDS: AN INTRODUCTION TO RECENT RESEARCH

PAUL A. DAVID
Stanford University

and SHANE GREENSTEIN
University of Illinois, Champaign-Urbana

This paper surveys the contributions that economists have made to understanding standards-setting processes and their consequences for industry structure and economic welfare. Standardization processes of four kinds are examined, namely: (1) market competition involving products embodying unsponsored standards, (2) market competition among sponsored (proprietary) standards, (3) agreements within voluntary standards-writing organizations, and (4) direct governmental promulgation. The major trajectories along which research has been moving are described and related to both the positive and the normative issues concerning compatibility standards that remain to be studied.

1. INTRODUCTION

Recent and ongoing advances in microelectronics and communications technologies have made the formerly esoteric subject of technical compatibility standards a familiar matter in our everyday experience, as well as a perennial topic for discussion in the industry press (see, e.g., Arnold (1985), Bartik (1985), Cropper (1980), Kass (1981), SJMN (1986), Teresko (1986), Witten (1983)). Today, many people—including academic economists—would readily acknowledge that considerable importance attaches to the existence and nature of technical standards for the encoding, storage, processing and transmission of information. Questions concerning compatibility and voluntary standards-setting have emerged as having central strategic significance in the development and marketing of computer operating systems and software (e.g. DOS and UNIX), value added data networks (e.g., TELENET and TYMNET), local area networks, high-definition television, and optical disks. The recent explosive

---

1 Preparation of this review was begun in connection with work supported by a grant from the National Science Foundation, Division of Information, Robotics and Intelligent Systems' Program on Information Impact. The authors are grateful to Joseph Farrell, W. Edward Steinmueller, and Peter Swann for comments and suggestions on earlier drafts. Richard McMasters provided excellent assistance with the final version of the bibliographic references and the text.

2 In addition, there were and remain many familiar, non-esoteric products for which the question of compatibility with other, complementary elements of a technological system is still significant for consumers' welfare and business success: agriculture equipment, appliances, audio equipment and program media, photographic films and lenses, typewriters (keyboards especially), automobile spare parts, and, more recently, color TV and VCRs.
growth in the use of facsimile (FAX) machines testifies in a positive way to the critical role played by compatibility standardization in expanding markets for network technologies.

A "standard" is to be understood, for the present purposes, as a set of technical specifications adhered to by a producer, either tacitly or as a result of a formal agreement. It is helpful to distinguish among several kinds of standards – reference, minimum quality, and interface or "compatibility" standards (David 1987c). Reference and minimum quality standards, when adhered to, provide signals that a given product conforms to the content and level of certain defined characteristics. Generically, these standards reduce the transactions costs of user evaluation. Interface standards assure the user that an intermediate product or component can be successfully incorporated in a larger system comprised of closely specified inputs and outputs. A product that conforms to an interface standard can serve as a subsystem within a larger system built from numerous components and subsystems that are provided by different suppliers, each of whom also conform to the same standard. One homely illustration of the latter is provided by the compatibility of "hi-fi" stereo sound system components.

Whereas "standards agreements" must be negotiated explicitly, "standards" more generally may arise in other ways. One is the widespread passive acceptance of a set of specifications that has been promulgated by a single agent acting unilaterally. Alternatively a standard can emerge spontaneously, through an undirected competitive process in which many individuals exercise their choices among a potentially wide array of alternatives.

"Standards", therefore, may be established by the widespread acceptance of any of the following: (a) "unsponsored" standards, these being sets of specifications that have no identified originator holding a proprietary interest, nor any subsequent sponsoring agency, but nevertheless exist in a well-documented form in the public domain; (b) "sponsored" standards, where one or more sponsoring entities holding a direct or indirect proprietary interest – suppliers or users, and private cooperative ventures into which such firms may enter – creates inducements for other firms to adopt particular sets of technical specifications; (c) standards agreements arrived at within, and published by voluntary standards-writing organizations; (d) mandated standards, which are promulgated by governmental agencies that have some regulatory authority. The first two of the foregoing outcomes emerge from market-mediated processes and are referred to generally as de facto standards. The latter pair usually issue from political ("committee") deliberations or administrative procedures which may be influenced by market processes without reflecting them in any simple way; both are sometimes tagged loosely as de jure, although the standards have the force of law behind them only in the last case.

Research on the economics of standards in information industries has been expanding rapidly during the 1980s, but the field remains young and in a quite fluid state. Economists have hardly settled on a standard terminology, much less converged on paradigmatic modes of theoretical analysis and empirical inquiry. It may seem a little premature, therefore, to undertake to codify and assess the state of knowledge regarding this new subject. On the other hand, some effort at stock-taking may be particularly useful while the program of research remains malleable – that is to say, before one or another approach to the subject of compatibility standards emerges as a de facto standard. Our aspirations in this review are modest: to survey the contributions that economists have made to understanding the standards-setting process and
its consequences for industry structure and economic welfare, and to relate these findings to fundamental issues that remain to be studied. By tracing the trajectories along which research has recently been moving, and locating these in reference to both positive and normative problem-areas — especially in information technology intensive industries, where issues of compatibility standards are of recognized practical importance, we may hope at least to provide a few guideposts for newcomers to this field.

The paper is arranged as follows. Sections 2 to 5 respectively examine the four types of standards-setting process identified in (a) - (d) above. This review summarizes advances both in the theoretical understanding of the dynamics of standardization processes and characterization of empirical patterns observable in the actual operation of market and non-market processes. In the course of reviewing the literature we will try to indicate where future empirical and theoretical work may go forward fruitfully, hand in hand.

2. UNSPONSORED STANDARDIZATION PROCESSES

There is now a substantial literature examining economic processes affecting the formation of “unsponsored” standards. Typically, these are situations where no agents have proprietary interests in any of the relevant standards and no firm or user is large enough to take account of the way their pricing and technology choice decisions may influence the actions of other vendors or users. Often these models serve as benchmarks that highlight the implications of underlying technical features of a product market — such as complementarities in demand, or network externalities in consumption; or of an industry’s production processes — such as technical interrelatedness (strict complementarities among inputs), or “learning” spillovers.

2.1. Increasing returns to adoption and “lock-in” by historical events

A number of analyses focus on issues that arise when standardization processes display increasing returns (to the marginal adopter) in the number of existing adopters. Increasing returns to adoption occur either through demand side externalities, network effects, or installed base effects, or, more generally, through cost reductions attributable to experience-based learning, or to the attainment of system scale economies. An over-riding theme in these analyses is that situations with increasing returns differ fundamentally in the dynamics of development, and in the comparative statics of their outcomes, from situations characterized by constant or decreasing returns in the extent of adoption. At the macro-level, dynamical systems of the former kind display “positive feedback”, whereas, in the latter two cases, respectively, there is no feedback (state independence) or negative feedback. As a general proposition, it is known that where there are strong positive feedback effects, the system will possess a multiplicity of stable attractors, or equilibria. Those points of equilibrium (in relative market shares held by the various technological standards) need not be situated at the extrema, where de facto standardization on one or another of the relevant alternatives is complete. Farrell and Saloner (1986a) present a deterministic model of technological competition in which there are positive “installed base effects”, and analyze the range of installed bases from which there are, and are not, multiple equilibria.

One paradigm for stochastic models of this kind has been provided by Arthur’s
(1983, 1988) analysis of unsponsored “competing technologies”, where each technology is subject to increasing returns that influence the choices of individual agents – the technologies are posited to be more attractive the more they are adopted, and it is assumed that agents who are differentiated in their inherent preferences among those alternatives enter the market in a random sequence and make irreversible (individual) commitments to one or another technology. (Also see David’s (1986a, 1987c) exposition, and Habermeier (1987) for a formulation emphasizing “learning by using”). If no countervailing factors serve to bound the increasing returns effects, the process eventually must “lock in” to a monopoly of the market by one technology; a point is reached after which every agent, regardless of inherent preferences, will select the same technology. Thus, it is quite possible for this kind of process to lead to de facto standardization on a technological alternative that is viewed, ex post, as Pareto-inferior to others that were available. Although every agent would have chosen to their own best advantage in the respective contexts in which they made their commitments, in the end everyone involved might prefer a different standard to the one at which they arrived. The stochastic process (described by a generalized Polya urn model) is non-ergodic or path dependent, in the sense that the “emergent standard” depends on the details of the historical sequence in which individual choices occurred, that is, on the path the process of adoption took. In this case, at the outset of the process, when installed base effects are very weak, the outcome is unpredictable; selection of one among the available equilibria is found to be especially sensitive to the chance events which dominate in the early adopters’ decisions – and so determine the shares of the different technologies in the growing installed base.

In some respects, these properties of unsponsored de facto standardization processes do not depend on the source of the positive feedback (i.e., whether the increasing resources arises from network effects rather than learning-by-doing externalities). Cabral (1987) has shown that in this regard processes of network externalities are analytically similar to learning processes, and results concerning the existence and stability of equilibrium can be established within a common general framework. In other respects, however, the source of the increasing returns shapes the analysis. Some propositions do depend crucially on the supposed inability of any one decision-maker to internalize the benefits associated with using a particular technology. Were one firm to internalize, say, all the learning benefits associated with one technology, as might occur when it controls that standard, then that (sponsoring) firm might price the product low at the start in order to encourage later adoptions.

Recent work has also demonstrated that the positive externalities at a local level can generate multiple-equilibria and non-ergodicity at a global level. David (1989) presents a heuristic model in which members of a finite population of agents each make (recurrent) technology selections at random intervals in time (corresponding to a stochastic replacement process), subject to the positive influence of the currently prevailing technology usage among a subset of “neighboring” agents. The latter belong to the local information field of the index agent. The stochastic process in this case can be described by a finite state continuous time Markov chain in which there are multiple absorbing states, each corresponding to a situation in which all agents’ technology choices become perfectly correlated. Complete standardization on one or

---

3 Models having the latter features have been explored by Hanson (1984), Farrell and Saloner (1986), Katz and Shapiro (1986), and, more generally, David and Olsen (1986, 1988), and Bresnahan and Trajtenberg (1989). See Section 3, below.
to be a certainty, and once the system arrives at any of those absorbing states, will persist indefinitely ("lock-in").

Another approach to modeling decentralized coordination processes driven by localized positive feedbacks builds on the extensive literature in mathematical physics devoted to dynamic Ising models of ferromagnetism; and to the general class of dynamical systems in which particles are subject to spontaneous random reorientations that are influenced by "additive" (rather than "cancellative") local interactions. In this vein, Steven Durlauf (1998) develops a dynamic model in which complementarities are local in the sense that the stochastic choices made by firms are positively influenced by the orientation of firms that are spatially "close" to that decision unit. Durlauf's formulation analyzes sectoral interdependence effects determining the time-series movements of the microeconomy between high and low levels of production. But, it seems that the model could be transformed to represent the interdependence among designers of system sub-components, whose payoffs tend to rise if they achieve compatibility with others' designs.

In many situations where coordination and compatibility) matters because there are complementarities, there seems to be a pie justifiably for presuming that some forms of "network effects" - either in demand or in supply - cannot be fully internalized by any one among the participants. Intertemporal externalities are thought to be difficult to contract around; future generations of standards users, for example, may not yet have representative interests in the present market place, much less know what features they will desire in their product standards. Cowan (1988), in analyzing a model combining a multi-arm bandit problem with positive feedback (dependent upon use), has shown how the inability to contract over time may lead present generations to investigate a technological alternative that offer a different technology that potentially offered greater benefits to later users. It is an essential assumption of Cowan's model that later adopters cannot contract with early adopters in order to change the decisions of the early adopters. Similar issues arise in the models of Farrell and Saloner (1986a, 1986b, Katz and Shapiro (1986), which will be examined below.

Expectations will not automatically solve these intertemporal problems. Arthur (1988), and David (1987c) point out that expectations may amplify tendencies toward lock-in, particularly if adopters' expectations are based on projections of the likely outcomes in a competition between technologies. Thus, even a small lead by one or another alternative becomes important, not only because of network benefits but also because of the nature of the lock-in, as well as the competitive dynamics. The conceptual paradigm underlying the models of increasing returns roughly corresponds with several empirical cases where crucial early events shaped the adop-
tion and development of standards over time. David (1985, 1986b) showed how the interaction of uncoordinated decisions by early typists and their employers, typing schools, and typewriter manufacturers, resulted in the adoption of the QWERTY keyboard layout. This case of unsponsored standardization is interesting because this keyboard arrangement—which has been carried over for use with computers—is thought to be ergonomically and in other ways inferior to several alternatives that have been available for some time. The Dvorak Simplified Keyboard (DSK) is probably best known among these, and was made an option on computers, but to date it has not been successful in inducing a switch from the QWERTY standard. Similarly, Cowan (1988) has analyzed how a series of circumstances affecting America's policy regarding commercial uses of atomic energy—such as the U.S. Navy's nuclear submarine program, the politics of the National Security Council, and the behavior of Admiral Rickover—helped to "lock in" the electric utility industry to light-water nuclear reactor technologies, despite the engineering superiority of gas-cooled reactors. And Richard Rosenbloom's research on the development of VCR standards has shown that the technical superiority of the more compact Sony Betamax format notwithstanding, the eventual dominance of the VHS format as the industry standard emerged from the interplay of adventitious and seemingly unrelated background conditions and events. A somewhat different, yet analytically related observation is made by David (1987a), and David and Bunn (1988), regarding the emergence of polyphase AC as the standard for electrical supply systems in the U.S. at the end of the last century. Though an exact correspondence between historical events and abstract models should not be demanded, it may be appropriate to ask whether the theoretical representation of "lock-in" as an indefinitely persisting equilibrium may not presume a stricter form of irreversibility than is warranted by the realities of technological and economic obsolescence. As a practical matter it may not be possible to distinguish between the indefinite and the merely prolonged (but finite) persistence of a configuration when, in either case, the system is most likely to be disrupted by external forces. Furthermore, analytical models of situations where standards for system components are in constant flux—as they are updated and revised by new market suppliers and customers with new demands—have yet to be developed, which is none too surprising.

4 Liebowitz and Margolis (1989) question the claim that the DSK (developed in the mid 1930s) was a superior alternative; and, presumably, also that the Ideal Keyboard—a late nineteenth century contemporary of QWERTY—also would have been preferable. They argue that this casts doubt on David's (1985) characterization of the episode. However, in their discussion they start from the premise that if there were a more cost effective keyboard than QWERTY, some firm would find it possible to make a profit by introducing it, which assumes the conclusion; and they fail to distinguish between efficiency ex ante and ex post—in the sense described below, in section 2.3.

5 Prominent among these were (a) the legacy of prior relations between Sony and the Hitachi Corporation, (b) the incidental ability of the bulky VHS cassette initially to carry a tape with a longer playing time, and (c) the unanticipated introduction of pre-recorded movies on video cassettes. See Rosenbloom, summarized in David and Greenstein (1989).

6 Although engineering evidence indicates that alternating current historically was superior to direct current for extensive power networks, due to lower transmission costs, DC had many other advantages as well as an initial lead in "installed base". In this case also it was extraneous and transient factors, such as the appearance of an AC-DC converter technology, and the personal inclinations and financial situations of key industry leaders—Edison and Westinghouse in particular—during the late 1880s, which determined the victory of AC in the so-called "Battle of the Systems".
in view of the rather intractable problems they appear to pose. It is not yet clear how the notion of “lock-in” and related ideas should be modified to account for a world in which once-stable points of equilibrium may suddenly be rendered unstable by endogenous processes. Are there degrees of “lock-in”? Are some features of a de facto standard more mutable than others, even though the original choice of an architecture constrains the ways in which features can be subsequently modified?

2.2. Bandwagons and coordination problems

A closely related line of research has focused on the role of information or communications, or the lack thereof, in permitting beneficial coordination to be achieved through the emergence of a de facto standard. Analyses in this vein almost always examine phenomena – such as the existence of multiple coordination equilibria – that are traceable to the presence of increasing returns or network externalities (see Cabral (1988)), and highlight the links between the sources of coordination problems and their solutions. Makowski and Adler’s (1988) discussion calls attention to the point that network externalities per se are not a necessary condition for the existence of coordination problems; they emphasize, instead, the role of complementarities in production or consumption, and the absence of organizational or institutional mechanisms that would coordinate the actions of buyers.

Some of the pioneering analyses of technology adoption decisions subject to network externalities (Rohlfs (1974), Oren and Smith (1981), Katz and Shapiro (1985), Farrell and Saloner (1985b)) noted that coordinating the launch of a new product or standard can present special difficulties. The canonical problem is this: if a large majority of the buyers use a particular technology or standard, then everyone benefits enough to justify the private costs of adopting that standard. However, if only a fraction of the market adopts a product, then for many users with low private gains from the new standards (in the absence of network effects) the total benefits may be insufficient to justify adoption of the standard. Hence, the two polar outcomes both are possible equilibria – virtually everyone adopts the new standard, or virtually nobody does. How will either outcome be “selected”?

The obvious intuition suggests that if a majority do adopt a new standard, so will the rest, and if a threshold number of adopters is never reached, swaying the rest of the market will prove impossible (see Granovetter and Soong (1986)). Farrell and Saloner (1985b) provide an insightful analysis of this coordination problem. They show that if all users would be better off with the new standard, then they will all switch to the new standard if each decision-maker has perfect information about the preferences of the others. Essentially, each decision-maker anticipates everyone else’s switching decision, and thus one decision maker after another switches, in a sequence from those with the largest private gains to those with the largest network gains. This dynamic process for overcoming coordination problems is sometimes referred to as a “bandwagon”, since as a standard gains adherents it becomes more attractive for others to climb aboard (see Farrell and Saloner (1986a, 1986b), Katz and Shapiro (1986)).

This full information case is also compared by Farrell and Saloner to situations in which agents are uncertain about one another’s preferences, but are aware that individual’s preferences differ in regard to the competing standards. As is common in models with positive feedback, it is found that early decision-makers, by committing
to one standard, can influence the economic returns to the choices facing later actors, and so can exert strong leverage over the process. Most interesting of all, uncertainty about others’ preferences can lead to either excess inertia or excess momentum in switching to the preferred standard. “Excess inertia” – according to Farrell and Saloner (1985b) – arises when no user is sure that she will be followed in a switch to a new standard, and is unwilling to risk committing to a switch without more assurance. Even if everyone would be better off if they switched, when preferences are “lukewarm” there is no incentive for any firm to start a bandwagon rolling. “Excess momentum”, on the other hand, arises when an early adopter so influences the returns to a later decision-maker that the second-mover also switches, even where the sum total of net benefits is higher when no switch takes place. Postrel (1988) has shown just how complex the dynamics can become in a model of multiple decision-makers. He finds that excess inertia and excess momentum can arise, both at the start of standardization process and at various points thereafter, conditional on a certain number of decision-makers having already adopted.

Improvements in communication among the agents need not eliminate the problems of momentum or inertia. Farrell and Saloner (1986a, 1986b) show that perfect communication about everyone’s preferences will solve the coordination problems when everyone would be better off with the same standard. However, when preferences differ, communication will be engaged in by rational actors with strategic motives, taking into account its effects on later users. Thus, better communications can reinforce tendencies towards excess inertia or excess momentum, rather than diminishing their strength.

Besen and Johnson’s (1986) illuminating empirical studies of standardization processes in seven cases involving telecommunication technologies, accords special attention to coordination problems. They extract the following general propositions from their case histories: (1) differences in the preferences of firms and users can interfere with market coordination in the collective choice of a standard; (2) lack of knowledge about others’ preferences, as expected, interferes with coordination; (3) strong demand for a product will speed the process and force fence-sitters to make decisions; (4) marketing strategies pursued by firms whose interests diverge from those of the majority, such as promotional pricing, advertising and other modes of standards-sponsorship (see below), can undermine voluntary standardization agreements. In another study, based on events in the micro-processor market, Swann (1985, 1987) considered a coordination mechanism that involved the behavior of chip designers. He observed that a variety of designs are put forward when a technology is still young, but subsequent imitators narrow the varieties available (while increasing the supply) by choosing to imitate only the most popularly used designs.

Future research should reexamine the historical experience of general standardization movements (e.g. those in the 1920s) and the market and non-market institutions that facilitate solutions to coordination problems. For example, Puffert (1989) discusses the circumstances that led to more than a thousand miles of remaining “non-standard” southern railway track gauge (width: 5’ 0”) being switched on two days in 1886 to a gauge compatible with the “standard” railway gauge (4’ 8 1/2”) used east of the Rocky Mountains. No governmental intervention was required by this undertaking.

---

1Of course, if the returns received by early adopters depend upon the choices that subsequent adopters will make, initial expectations about the distribution of preferences among later adopters also could powerfully influence the eventual outcome. See Haltiwanger and Waldman (1987), Habermeier (1987).
but the existence of an active cartel organization - in the Southern Railway and Steamship Association, appears to have been an important mechanism for coordination. Other, analogous episodes await careful comparisons. 8

2.3. Social optimality of standards selection and timing

In self-reinforcing processes (i.e., dynamical systems characterized by increasing returns, learning, and other sources of positive feedback) social optimality is problematic. As is generally the case in discussing decision-making under uncertainty, it is useful to distinguish between the ex post social optimality of the outcome (and of the sequence of choices leading to it), and the optimality of the choice process ex ante. Even with optimal decision strategies, low probability events can reinforce a course of action (through subsequent, conditional choices) that issues ultimately in a regrettable outcome - one that leaves the decision-maker(s) worse off than they would have been had the realization of the stochastic process conformed more closely to expectations.

But when the dynamic sequence involves externalities, so that agents do not consider the consequences of decisions made at the moment for the actions of other agents who will decide later, the process as a whole will not be socially efficient in an ex ante sense. By chance, of course, the actual outcome could be a globally efficient one, ex post. Perfectly deterministic systems with dynamic increasing returns do not guarantee the global efficiency of decentralized, de facto standardization processes. In stochastic systems with the same properties, the social efficiency of the sequence (or "path") is not assured either: it is possible that after some point all the agents would regret that a standard had gained so many followers, and yet none would find it rational individually to adopt another standard.

Cowan's (1988a) analysis of a two-armed bandit problem subject to increasing returns illustrates these points and the distinctions on which they rest. The model supposes that a technology choice, between two alternatives, must be made under uncertainty as to the actual payoffs, but that each trial results in some endogenous improvement of the technologies, enhancing their expected payoffs. A central policy maker, by taking into account the informational gains from conducting a parallel experiment for some period, could follow a strategy that yielded a sequence of choices that was socially optimal ex ante - after some period of information acquisition one of the two technologies always would be used. If the choice process was decentralized, however, the different agents would not take into account the effects of their conditional private optimization actions upon the information available to future decision-makers, so that an externality would exist. Notice that in this analysis better information is not assumed to be available to the central (governmental) decision-maker. When there are no externalities to distort the decision process it is only with a sense of backward-looking regret that one can say that some choice in the sequence turned out to be sub-optimal ex post. In other words, there would have been a run of bad luck that led to the selection of a technology that would have been dominated

8 Market episodes that may serves as useful comparisons include: (1) de facto coordination on VideoCipher, after Home Box Office chose it as the protocol for scrambling signals (Besen and Johnson 1986); (2) the slow initial coordination on one standard in the AM stereo market (Besen and Johnson 1986); (3) the world wide incompatibility of television systems (NTSC, PAL, and SECAM) (Crane 1979); (4) or the initially slow and then rapid convergence to a standard in the VCR market (Rosenbloom in David and Greenstein (1989)).
by the alternative, had a different realization occurred which led the alternative to be
tried and improved through usage. With hindsight one could calculate whether a
technology that offered initially lower benefits would have yielded greater long-run
gains, so that all the market participants would have been made better off were it to
have been selected as the standard.9

Other issues arise when evaluating the social optimality of the timing of standardi-
dization. That a market has chosen a standard, even "the best" among those avail-
able, ex post, does not imply that the timing of the process was socially optimal — in
either the ex post or the ex ante sense. So long as there are intertemporal externalities,
individuals' incentives will not reflect the indirect costs incurred the benefits accruing
to others, from the absence or presence of a consensus among suppliers and/or users
concerning technical characteristics of complementary products and production
methods. Standardization's consequences for innovation, and for market structures,
should be considered in this connection.

To be more specific, for example, early standardization of products may encourage
invention in complementary technologies and organization, and it may promote
subsequent incremental innovations designed to perfect the original technology. Such
developments are likely to be socially welfare-enhancing. On the other hand, de facto
standardization may prematurely close off basic exploration of technological oppor-
tunities in a wide area, "defining" the market for a broad class of production in
specific terms that discourage further investment in non-incremental innovations.
(This argument is further developed by David (1987c), and David and Bunn (1988).)
While the standard selected may appear "best" among the existing alternatives, the
latter may present an unnecessarily narrow range of options.

Analogously, standardization can affect market structure, by reducing entry costs
and risk for new firms on the supply side of the market for components. But if this
leads to intensified price competition and reduced profits, the ability of the industry
to sustain investments in improving the quality of components may suffer.

Issues such as the ones just delineated are especially deserving of study in the setting
of industries that are experiencing rapid and continuous technical evolution, as is the
case in so many markets for information technology products and services. Unfor-
tunately, the speculations reviewed here remain largely unsubstantiated by systematic
empirical (or for that matter analytical) findings.10

3. SPONSORED STANDARDIZATION PROCESSES

Sponsored standardization processes differ sharply from unsponsored processes. In
the former proprietary control can create incentives for firms to manipulate technical
standards so as to make their goods compatible with complementary components or

---

9 After the system has locked in to one option, it is the nature of the privately optimal strategies for the
sequential decision ("bandit") problem that agents will never wish to explore the alternative since no other
option could ever be as good. Hence, "regret" by decision-makers in the model is, strictly speaking, not
possible. This is comparable to being married for twenty years to one person and then meeting another and
reckoning that twenty years with the new person would now leave you better off, had you met twenty years
ago. However, given the investment already made in the older relationship and that would have to be made
in a new one, it does not pay to "remarry".

10 See, however, the set of case studies by Putnam, Hayes and Bartlett (1982) of the effects upon
innovative activity of industry standards, which includes one instance of an interface standard.
substitute systems, as well as to engage in strategic price-setting. As a consequence, the sponsoring agents in the process take actions that anticipate the reaction of rivals, because they know their actions will affect the returns to adoption of the alternative standards. However, although analysis of sponsored processes must account for strategic behavior, as has been seen for the class of situations where a large number of firms are making choices among unsponsored alternative interface standards, rivalries involving sponsored standards often yield problematic outcomes that would not be expected in more conventional forms of industrial competition where constant or decreasing costs of production are the rule.

The practical importance of sponsored standardization processes derives from the influence of the ownership of assets. Initial asset ownership conditions — often resulting from previous episodes of market competition — can influence the evolution of subsequent standards, because initial asset ownership gives an advantage in the design and production of related components of a developing system. Such advantages, in turn, determine the abilities of particular firms to take on leadership roles in the design of interrelated system of components that become de facto standards. Because the situations can be complex, so too are the analytical issues. Yet, market rivalries involving sponsored standards have become increasingly common, especially in the areas of computer and telecommunications technologies.

3.1. Strategic behavior and competing standards

Some studies have compared the dynamics and outcomes of sponsored cases with those of unsponsored cases. An important issue here is whether pricing behavior by a standard’s sponsor can mitigate the effects of some of the intertemporal externalities that typify competitions among unsponsored standards. Hanson (1984) analyzed pricing behavior of duopolists, each of whom controlled a technology-product (incompatible with the other’s) that was subject to dynamic increasing returns. Using a modification of Arthur’s (1983) model, Hanson postulated that the use-value of each of the technology-products for a potential customer was an increasing function of their respective installed bases. Under the assumed cost and demand conditions, it is shown that the market eventually will “lock in” to only one of the products, giving its sponsor monopoly power. As a consequence, profit-maximizing firms will price aggressively in early periods in order to broaden the installed base and win the prize of “monopoly rents”. Despite this foresighted competition following so-called “penetration pricing” strategies, the duopoly market has a positive probability of “lock-in” to a technological standard that is socially sub-optimal, in the ex post sense. The outcome of the rivalry remains sensitive to early random events that may build up the installed base for one, rather than the other duopolist.

The effects of intertemporal pricing by a sponsor was investigated also by Katz and Shapiro (1986), who showed that strategic pricing by duopolists can help bring forward in time some of the eventual cost advantages of a dominant technology — because its sponsor lowers the price today in an effort to secure a larger installed base. Yet, in this analysis, too, Katz and Shapiro find that the socially optimal technology — the one that would yield a monopoly with lower costs — does not necessarily prevail in the competition. In a somewhat different formulation of essentially the same problem, Farrell and Saloner (1986a) allowed for dynamic strategies with continuous variation in pricing. They show how a sponsor of a technology who holds a larger installed base can price the product to forestall a rival, preventing the establishment
of a significantly large installed base to tempt customers to switch. This analysis draws attention to the point that a "window of opportunity" may exist before the technology with the largest installed base gains a large and insurmountable advantage (analogous to a "lock-in"), and that a sponsor need only pursue "promotional" (or "penetration") pricing until that window closes.

A few case studies have investigated the appropriateness of applying this conceptual framework in concrete historical settings. Saloner (1990) analyzes the consequences of changing from competition among proprietary systems to competition with a non-proprietary and "open system". For the case of UNIX-based computer systems, he is able to show rather neatly how the existing body of theory indicates the factors that are likely to be important in determining who will win and who will lose under different types of competitive environments. To cite another example, Besen and Johnson's (1986) account of the early competition between variants of AM stereo describes behavior resembling the sort of promotional pricing predicted by models of sponsored standards rivalries.

Economic theorists have analyzed several strategies apart from aggressive pricing that sponsors of standards may pursue. Farrell and Gallini (1988) observe that a monopolist may be willing to sacrifice full short-run exploitation of its market power - by inviting another firm to second-source - as a means of convincing buyers to use the technology. This is similar to the often-noted returns associated with the strategy of giving away a technology (or licensing a patented process or product at a nominal fee to many firms) in order to establish it as a standard (see Shoch in David and Greenstein (1989)). The latter strategy resembles promotional pricing where the price is initially set at or close to zero. A sponsoring firm may benefit also by granting licenses on easy terms, when it has an interest in selling complementary products that use its interface. Notice that a positive network externality could accelerate the adoption process once an installed base is established, and thereby enhance benefits to the sponsoring firm. In addition, as demonstrated by the dynamics of the UNIX case (see Saloner (1990)), technology sponsors can enter coalitions with potential rivals in order to generate a bandwagon in support of their standard.

3.2. When compatibility itself is endogenous

While one branch of the literature treats market competition between two product designs embodying (sponsored) standards that are technically incompatible, a related line of research relaxes the assumption that the degree of compatibility is exogenously specified. Once the possibility is recognized of designing compatible interfaces \textit{ab initio}, or of redesigning (at some cost) so as to remove incompatibilities, it becomes important to consider the incentives a sponsoring firm may have to render its products interoperable with those of rival firms in the same market, or with (potentially) complementary products.

A central insight which examination of this question has yielded is that initial conditions can matter a great deal in determining firms' strategies when compatibility is a design option. This is because asymmetries in market position give firms who sponsor alternative standards quite different payoffs from providing for "interoperability" (or realized technical complementarity) with competitors' products, especially when network externalities are important to the consumers. Katz and Shapiro's (1985) analysis clearly illuminated this point by considering endogenous pricing behavior in the presence of network externalities. The firm holding the larger
market share (initially) enjoys a measure of market power due to the high valuation that customers will place on a product that gives them access to the more extensive network. Permitting rival vendors to offer their customers access to that network would strengthen their competitive position and tend to erode the larger firm's profits.

For the firm with the larger network or installed base, therefore, the relatively small gain in the value of the product achieved by establishing compatibility with a rival's network does not necessarily counterbalance the decrease in rents caused by the enhanced attractiveness of a smaller rival's product. Firms with the smaller network or installed base, on the other hand, stand to gain from the increased willingness to pay for their product if it becomes compatible with the dominant network.

Treating compatibility as a continuous rather than a discontinuous variable does not change the above result, though it alters the appearance of the analysis. Duopolists may not have the appropriate incentives to remain compatible even when network externalities are present, because the degree of compatibility increases the ease with which some customers could substitute between the two firms' products. Berg (1985, 1988) analyzed this issue in a model where duopolists could be "partially" compatible with one another, a condition he represented by their closeness within product characteristics space. Firms traded off competition with compatibility and did not become fully compatible, though it was socially optimal for them to do so. The foregoing line of analysis may bear still more fruit. Neither Berg nor Katz and Shapiro (1985) linked the network externality to any specific structural features of their model. Thus, they left open questions of whether different sources of network externalities provide different incentives to remain compatible despite competitive pressures.

Moreover, neither model places these decisions in the context of dynamic market growth and development, where product (or system) life cycle issues also influence decision making.

Besen and Saloner (1988) took a very broad approach to these questions in their classification of market standards-setting processes (similar to that in Besen and Johnson (1986)). They argue that it is relevant to compare a firm's "vested interest" in a standard with the "degree of agreement" among all firms in promoting the universal adoption of a standard. In cases where vested interests are low, as when firms have invested little in the development of products, then two outcomes are possible. If the degree of agreement for a standard is high, then all firms have a high incentive to participate and standardization reduces to a coordination problem, analyzed above in the unsponsored case. If the degree of agreement is low, standardization may be slow to occur without government intervention because no individual agent has the incentive to develop a standard. When vested interests are strong, then two other outcomes are possible. First, if preferences for standardization are intense, then those with something at stake try their best to win a market battle of systems. The Betamax/Sony contest was an example of this type. Second, if preferences for standardization are weak, then Besen and Saloner expect the outcome to depend on many factors such as the existence of dominant firms, the formation of coalitions of firms, and the relative ability of government to coordinate actions. This approach nicely classifies several cases and has the promise of clarifying the links between characteristics of markets and the propensity to standardize – once tighter theoretical links can be drawn between the variety of observed market structures, forms of vested interest, and degrees of standardization.

Further insights into the incentives to remain compatible are provided by the literature about competition when there are costs to the buyer from switching between
alternative suppliers of essentially the same product (von Weizsacker (1983), Klemperer (1987a, 1987b), Farrell (1987), Farrell and Shapiro (1988, 1989)). These analyses can be reinterpreted as competition between vendors of incompatible systems. The question that arises in regard to standardization is this: if the vendors could choose (ex ante) to be compatible and abolish all switching costs, would they do so? Are profits higher with switching costs than without, everything else being the same? The answer is "maybe". Once a firm has sold a product, obviously, it wants to have the buyer face a higher cost of switching to a substitute product (sold by another firm). But, if buyers can perfectly anticipate all future switching costs (and the resulting market power of the incumbent vendor when replacement purchases are made), then, in the competition among vendors to gain the "incumbency advantage" it is possible that the present value of the future rents will be dissipated. The question remains open for further investigation, however, because results of the foregoing sort are sensitive to specifications of the model, particularly in regard to the extent of entry of new system and component suppliers, the enforceability of contracts, and the ability of incumbents to commit to strategies (See Farrell (1987) for a concise summary of some of these modelling issues.) Moreover, the whole structure rests on an empirical presumption that purchasing organizations and vendor organizations are sufficiently monolithic that they can formulate and implement consistent intertemporal strategies.\(^{11}\)

A number of theoretical studies suggest that in markets for interdependent (complementary) products, firms may be induced to design for compatibility even in the absence of network externalities in consumption. Matutes and Regibeau (1987, 1988) have pointed out that if firms make their products compatible, then consumers are able to "mix and match" components from different sources to create a customized system, much as sophisticated purchasers of stereo sound systems do today, and that this freedom increases the willingness to pay for the constituent components. Where incompatibility is maintained, customers are obliged to choose among essentially pre-packaged systems, or a narrower range of systems that can be constructed by permuting the array of components available from a single vendor, as is the case in the market for cameras and lenses. Matutes and Regibeau have shown that in a duopolistic market setting where the rival suppliers opt for compatibility, product prices and social welfare will be higher. But, total consumer surplus could be reduced, as some consumers are made better off while others are hurt.

The results obtained by Matutes and Regibeau (1988) regarding the value-enhancing effects of achieving compatibility among complementary commodities have been extended by Economides (1989a) to the case of an oligopoly that faces consumers whose preferences take a less restrictive, specialized form. For a given number of firms, it is found that product prices and vendor profits are higher under a regime of compatibility, essentially because demand for any one component is more elastic under the regime of incompatibility; intuitively, this follows from the fact that the demand for a pair of components from the same supplier will be linked under conditions in which different suppliers products are incompatible. Economides argues

\(^{11}\)Some considerable doubts are cast on the latter proposition by Greenstein's (1988) study of U. S. Federal Government agencies' behavior in the market for mainframe computers, which examined the long period during which (for reasons not all of which were peculiar to a governmental bureaucracy) hardware procurement decisions were effectively insulated from consideration of software - and personnel - related switching costs.
that free entry will not substantially disturb the foregoing comparative price results, but that under the regime of full compatibility there will be a socially inefficient degree of variety in the produce space of the components. In another paper, Economides (1989b) considers the more general situation where partial (rather than all or nothing) compatibility can be designed, and finds that full compatibility is generally socially advantageous. The exceptions are cases where the scale of demand for “hybrid” systems is so low that gains to compatibility do not make up for the increased competition that is unleashed by a regime of full compatibility with free entry.

3.3. Gateways as alternatives to ex ante compatibility standards

The contributions reviewed in the preceding sections implicitly assume that technical incompatibilities, once committed to in the designs of durable products and production facilities, will persist throughout the service lives of those assets. Such incompatibilities cause static inefficiencies. On one side, use-value of the existing stocks of production system components (i.e., of installed hardware and software) would be raised to the extent that compatibility widened the domain of complementarity. On the other side, the (social) costs of providing the same flow of services would be reduced to the extent that compatibility rendered a broader range of components more completely substitutable for one another. Consequently, the availability of converters, translators, emulators, and other “gateway technologies” that achieve compatibility ex post, serve to reduce the social costs of failures to standardize ex ante (as has been pointed out by Braunstein and White (1985), and David (1986a)). Such devices can be provided as an option by a vendor of a product that otherwise would remain incompatible with components supplied from other sources. Or, third parties may enter the market as specialized vendors of converters or of “gateway services”.

As an historical case in point, David and Bunn (1987, 1988) examined the circumstances and consequences of the development of the rotary converter which permitted conversion between alternating and direct current (and vice versa) – a devise introduced commercially by an independent inventor, and subsequently produced and marketed by the Westinghouse and General Electric companies. Generalizing from this instance, in which the converter facilitated the integration of pre-existing (local) DC electricity networks into larger systems based on AC, David and Bunn argue that in addition to the short-run resource saving effects, the evolution of a network technology can be strongly influenced by the availability of a gateway innovation, and the timing of its appearance. Although the converter itself was symmetric, its impact upon the competition between the DC and AC standards was non-neutral. This case, however, is one in which neither of the developers of the contending technologies (Edison and Westinghouse) initially foresaw the possibility of an AC induction motor, on which the converter was based. What happens to the balance between the social benefits and the costs of providing this “fix” for the failure to standardize at the outset, when the availability of a gateway technology is anticipated? There are two aspects to this question: the effects of foresight on the costs of performing the necessary conversions, and the influence that knowledge of this possibility will have on the demand for converters rather than for compatibility standards.

The cost of providing converters may be influenced by the variety of distinct technical systems that need to be made interoperable, as well as by the number of dimensions in which alternative technological standards diverge, as David (1987c) points out. One implication of this has been explored in Economides” (1989b)
analysis of a model that allows rival sponsors to affect the price at which customers can achieve compatibility, by varying their proprietary designs and thereby affecting the costs faced by a third party supplier of adapters or converters. Firms that stand to gain from having their products become joined to a larger network would not seek to forestall the entry of a gateway technology, whereas a dominant sponsor might design products with a view to raising conversion costs. Economides (1989b) suggests that duopolists have an incentive to reduce the costs of achieving *ex post* compatibility by the addition of a converter. But, in qualification of this we have Farrell and Saloner’s (1989) theoretical analysis of the economics of converters, which concludes (albeit from a somewhat different model) that duopoly pricing policies lead to less *ex ante* standardization and greater reliance on converters than competition or monopoly price-setting, and that the duopolist that holds the dominant network wants converters to be privately costly to the users (but not technically inefficient).

The latter result suggests that the (foresighted) manipulation of technological designs may curtail the resource savings that are available from investment in the development of converters. Farrell and Saloner (1989) bring out a further point which carries the same implications, and which likewise follows from recognition of the ability of agents to anticipate the possible *ex post* development of technical gateways through the walls created by failure to achieve *ex ante* interface standardization. Although converters reduce the social cost of a failure to standardize, because they also reduce users’ private costs of indulging their diverse tastes, they make such failures more likely. When each of two technological systems characterized by positive network externalities, and (two-way) converters, are supplied under conditions of perfect competition, Farrell and Saloner (1989) find that if there is an equilibrium in which converters are used, “too many” will be deployed.\(^\text{12}\)

Although it might be supposed that this inefficiency would be eliminated, or at least reduced, if all of the technologies were supplied by a monopoly, the opposite turns out to be the case when the monopolist cannot price discriminate, for the latter is finds it profitable to price in such a way that the marginal buyer on each variant technology uses a converter to derive a large benefit independent of which network their equipment is directly compatible with. Thus, according to Farrell and Saloner’s (1989) analysis, the pricing of alternative types of network terminals and converters by a monopolist that does not discriminate among users (with different tastes) actually exacerbates the problem of socially excessive use of converters – pushing the latter above the level that would be found were the network technologies to be competitively supplied. David and Steinmueller (1990) discuss the implications this line of reasoning carries in the context of the internal policies that organizations adopt with regard to procurement and allocation of computers and peripheral equipment (e.g., printers) that can be integrated within local area networks through the use of gateway devices.

Many interesting issues in this area remain unexamined. In view of the multiplicity of pricing and promotional strategies typically available to firms, there is a need to arrive at a better understanding of the economic incentives that lead products to be

---

\(^{12}\) Briefly stated, the reason is that because the conversion process is not perfect, users who purchase a product designed for perfect compatibility with the dominant network confer a greater network externality on others, in comparison with those who buy the minority network product and a converter. Since the individual user does not take account of this, the externality reflects itself in a market failure: the outcome is that there is less designed compatibility than there would be in the absence of the converter option, and excess reliance upon the (imperfect) converter technology.
designed so that they are compatible, or incompatible with the present and future generations of systems of products. It is also not clear what conceptual framework is most appropriate for many empirical settings. For example, we do not know what structural features of markets empirically determine observed differences across product markets in the extent to which firms readily make their products compatible with others (e.g. contrast stereos and cameras). Despite the advances represented by the explicit incorporation of gateway or converter options in formal models, the framework of analysis remains unrealistically static, telescoping all the effects within a single period, and possibly exaggerating the adverse consequences of making converters available by imagining only situations in which they are excessively deployed because, in a sense, they are made available “too soon.”

3.4. Market structure and systems of interrelated components

Structural features of markets, such as the number of firms, and the degree of vertical integration typical of suppliers, can affect the success of efforts at interface standardization in industries where communication network externalities are important. (For a case study that takes this approach, see Sirbu (1986).) In these industries, when either buyers or sellers are initially capable of unilaterally acting to assure technical compatibility among the sub-systems of a larger system, diversity of designs and interfaces need not be dysfunctional, and possibly may be profitable for the vendors. On this view, universal (voluntary) adoption of formal standards is likely to emerge only where there is fragmentation among the buyers or the sellers of technically interrelated components. Where market participation and firm sizes are essentially pre-determined, as often is the case in the short run, this framework for analysis can provide useful insights.

For long-run analysis, however, such an approach can be confusing, unless clear distinctions are drawn between features of market structure that properly are treated as initial conditions, and those that are jointly determined with the degree of standardization of product specifications. In other words, both the extent of standardization and the industry’s evolving market structure may be viewed as endogenous variables. A distinction should be made between the initial distribution of ownership of assets (including intellectual property) for the interrelated parts of a system, and the subsequent pattern of firms’ participation in the designing, manufacturing, and marketing of those components. The degree of integration that is observed to persist within the industry needs to be analyzed as a (temporary) equilibrium result, which may reflect the influence of both cooperative standards-writing processes and strategic competition, rather than being treated as an exogenous structural parameter that determines individual firms’ policies vis-a-vis compatibility standards.

One set of insights into these questions has come from analyses of second-sourcing arrangements – the intentional development of second suppliers by firms that (at least momentarily) control a technology. Swann’s (1985, 1987) studies of the development of industry standards for microprocessor chips contend that buyers sought to guarantee multiple sources of supply and to increase the probability of future support for a particular technical design. This led vendors to take actions that encouraged agreement de facto on a few variant designs with multiple suppliers, rather than the persistence of many vendors, each sponsoring a different design. Farrell and Gallini (1988), and Shepard (1987) have clarified some of the strategies associated with this phenomenon. The first of these contributions shows how second sourcing can help
commit the vendor to not exploit monopoly power through raising prices once buyers are locked in – when switching costs are large. The second showed how buyers can use second sources to avoid the capacity limits of a single vendor. It would be valuable to link these insights with those developed earlier about pricing behavior early in the development of a product's life cycle, especially when systems evolve over time.

Several case studies of standardization processes have examined the subtle interactions of initial market structures, standardization and long-run market development. Bresnahan and Chopra (1990) contrast the development of standards for LANs in the office and the factory, emphasizing the inherent conflicts between buyers, who almost uniformly desire adherence to standards by all suppliers, and vendors, who would prefer to market proprietary systems that "lock in" buyers. How this issue was resolved in practice depended crucially on the initial market structure of vendors and buyers. In the case of factory and office LANs, it depended on whether any firm was large enough to coordinate the entire market, be it GM as a buyer, or IBM as a vendor. In contrast, Brock's (1975) account of the development of COBOL and ASCII standards for mainframe computers concluded that the divergent manufacturers' interests – especially the tensions between IBM and the other firms – prevented voluntary self-regulation from accomplishing a socially more efficient outcome.

Though few studies of standardization have done so, one also might argue that coordination on one standard is the fortuitous by-product of new innovation which embodies new standards and replaces obsolete technology – a possibility that is discussed briefly by Carlton and Klamer (1983). Thus, the emergence of standards could be analyzed in terms of the degree of coordination achieved between buyers and sellers as a consequence of vertical integration or close subcontracting relationships (or the lack thereof) for the purpose of supplying a complete, integrated system. Teece (1989) has partially pursued this line, arguing that vertical integration into the production of related components gives an added incentive to innovate. The reason is that an integrated vendor can better coordinate decisions related to the design and marketing of new and technically complex products, which typically are composed of many interrelated components.

It would be helpful in pursuing this line of analysis to know more about the importance of such considerations, including the impatience in waiting until the market converges on de facto standards, among the motives for expanding the scope of the firm into the production of technologically complementary. This requires understanding how the existence of a "full system supplier" influences the standards-setting process. For example, how did IBM's support of FORTRAN affect the acceptance of this widely used programming language? Are there incentives for a firm to integrate into multi-component production as a means to increase the likelihood of market acceptance of its proprietary technical designs as de facto industry "standards"?

Three other areas of empirical study suggest fertile research angles. The first, which few economists have pursued, concerns the interaction of the emergence of standards and the life-cycle of a product. Thompson (1954) analyzed the emergence of standards in the early automobile industry, arguing that standards facilitated industry growth by accelerating the use of interchangeable parts. In the early industry, standardization helped suppliers achieve economies of scale and it lowered the costs of repair and parts replacement. Clearly his argument depended on the timing of the introduction of standards into product development.

The second arises in economic analysis of the demand for ISDN (see, e.g., David
and Steinmueller (1990), Lehr (1989), and Lehr and Noll (1989)). Several observers have argued that the standard’s viability is closely linked to the commercial viability of the products in which it is imbedded, and therefore to the alternatives that are available to buyers. Many of the products are so young that the standard’s technical feasibility and economic value is much in doubt. As a consequence of this uncertainty, it is difficult to predict whether the technical requirements implicit in ISDN will limit competition on a network or enhance it.

The third emerges from the policy issues regarding regulation of the telephone network, where many problems are specific to this network technology. The general issue concerns access by several vendors to a network of interrelated components. What are the optimal points of “access” to the existing telephone network for competitive component suppliers? Which interfaces should best be left to independents (see Dunn (1987))?

There is still a lot to be learned about the factors that induce firms to market products incorporating proprietary standards when such decisions have consequences for long-run market structure. For example, if the rents from innovating in certain components markets of a system are not easily appropriable, how will this alter the firm’s incentives to design a proprietary system? In addition, if firms recognize that standardization of products is important for their own success, then they may seek vertical integration as a means to gain the advantages of size, or of becoming a full system supplier. Alternatively, vendors of fully assembled systems may find that non-proprietary industry-wide standards for component parts have the effect of lowering entry risks for parts suppliers, as well as lowering transactions costs in their own dealings with sub-contractors.

3.5. Dominant firms’ strategies regarding compatibility

There is a long-standing tradition of concern that large firms may exert a disproportionate control over the terms of market competition, by not only setting prices but manipulating product quality in ways that are privately profitable but not socially efficient. Inasmuch as compatibility is an aspect of product quality, it is hardly surprising that some contributors to the antitrust literature have alleged that a dominant firm has an incentive to manipulate the interface between system components over which it has established proprietary control, and other, potentially complementary system components. The ostensible purpose of such manipulation is to make it more costly for rival producers (or third party providers of converters, as we have seen) to achieve compatibility.13 The thrust of the recent theoretical work relating to the dynamics of technological rivalry under increasing returns conditions has simply expanded the a priori grounds for worries on this score. Even a firm that holds a transiently large market acquires important leverage over the dynamics of a market-mediated standardization process. As has been noted, the nature of bandwagon effects that produce “excess momentum”, as well as the strength of incentives to achieve coordination through negotiation, are sensitive to the relative shares of the market held by the competing firms.

Closely related to these issues are those concerned with the advantages dominant firms may acquire in markets for complementary products, even when there are no

13 For allegations of this kind directed against IBM, see, e.g. Adams and Brock (1982); and the rebuttal by Fisher, McGowan and Greenwood (1983).
strong network externalities. The situation envisaged here is one in which a firm faces competitors in its home market (e.g., where mainframe computer vendors face one another) and in the market for a complementary product (e.g., where vendors of mainframes and compatible tape drives face specialized producers of tape drives). Questions about the advantages obtainable through the “anti-competitive” manipulation of interfaces have arisen typically in the context of antitrust litigation in the telephone and computer industries, but they appear also in the market for home photography products and automobiles – where network externalities are thought to be hardly so significant.14

The circumstances in which systems developers may have an advantage over single component producers, and in which integrated system suppliers can manipulate interface standards for their private benefit, and to the disadvantage of rivals, have been examined in studies by Adams and Brock (1982), Braunstein and White (1985), Carlton and Klamer (1983), Fisher (1979), and Greenstein (1990). A related branch of the literature examines whether product innovation in systems of interrelated components is always beneficial or may be “predatory” in some sense (see Besen and Saloner (1988), Ordover, Sykes and Willig (1985), Ordover and Willig (1981), Stockdale (1979)). For present concerns, the central issue is whether a firm that monopolizes the production of a key component can enlarge the sphere of its market power in product space by manipulating interface standards in a way that effectively “bundles” its proprietary component technology with complementary components over which it holds no exclusive rights. This problem has been addressed explicitly by Bowman (1973), Greenstein (1990) and Whinston (1988). As a fine summary of most of the literature devoted to these issues has been provided in Besen and Saloner (1988), only a few comments on recent contributions will be offered here.

The typical theoretical treatment proceeds by assuming a dominant firm produces an essential component for a multi-component system. It is then supposed that one or another among the following strategies can be pursued: refusing to sell the primary good to a rival, selling only complete systems and not their components, selling both system components but setting high prices for components if purchased separately, “underpricing” components that compete with those sold by rivals, and “overpricing” components that are needed by rivals to provide complete systems (this follows Besen and Saloner (1988, p. 48)). Thus, exploiting control of an interface standard in many instances is equivalent to a strategy of “raising rivals” costs”, which has been analyzed in considerable detail by Krattenmaker and Salop (1986a, 1986b).

Farrell and Saloner (1986a) added several novel insights, drawn from their model of a product market situation where the installed base of the dominant incumbent is not large enough to guarantee “lock-in” of its standard, and a “window of opportun-

---

14See Besen and Saloner (1988) for a thorough review of the relevant issues as they pertain to telecommunications.


ity" therefore exists for the entry of an alternative standard. By pricing its product aggressively, as has been noted, the dominant incumbent may be able to deter entry for the duration of its vulnerability. Farrell and Saloner examined the advantages to be gained in this regard through the use of an alternative strategem: "compatible product pre-announcements" may forestall switching on the part of consumers who otherwise would be attracted by features of new entrants' technologies. If entry is successfully blocked in this fashion for a period long enough to allow the dominant firm's installed base to reach the critical mass, or "lock-in" point, the latter may delay actual delivery of the pre-announced product. Such behavior would give rise to the phenomenon of "ghost products", referred to derisively as "vaporware".

The allegation that technical standards are being used to extend monopoly power has arisen repeatedly in antitrust cases involving "leveraging", a concept that has been rather suspect among economists in the field (see, e.g., Bowman (1957) and Posner (1976)). Two recent contributions, however, have departed from the implicit assumptions that appear to have underlain the courts' generally critical stance towards the charge of "leveraging." Whinston (1988) examines the uses of product tying, or "bundling" - which here can be reinterpreted as manipulation of interface standards - to reduce the profitability of entry in a complementary product market. Abandoning the traditional assumption that components all are produced at constant marginal cost and sold under perfectly competitive conditions, his analysis supposes there are economies of scale in the production of the tied good and the market structure is oligopolistic. Under these conditions tying is found to decrease the sales of rival firms and, because it induces exit among the latter, constitutes a profitable strategy for the primary component producer. With suitable reinterpretation, this suggests a rational motive for retaining control of interface standards.

In a second break with conventional wisdom in this area, Greenstein (1990) discards the static framework within which formal analysis has been conducted, in favor of a model in which it takes time for imitators of complementary components to enter the market - because they have to redesign when interfaces with a primary component are altered. A primary component producer thus may have an incentive to manipulate interfaces, if the direct private costs of doing so are small and the change suffices to delay imitators and afford an interval in which to collect rents. Although the concept of leveraging in these contexts finds a measure of theoretical support, it needs to be used with considerable care as a basis for regulatory action. Artificially changing technologies certainly is a wasteful activity, but the economic welfare consequences for society remain ambiguous, because the creation of monopoly power may permits price discrimination that can improve the welfare of some groups of users while shifting the social costs to others.

Economists have yet to develop fully a normative analysis of the role of dominant firms in nascent industries, where standards remain under-developed. In the presence of a demand side externality, or of increasing returns, will a sponsored market process result in the "correct" amount of standardization and the optimal amount of variety? Will sponsored standards be adopted too slowly, thereby foregoing the benefits of compatibility and the expansion of the markets for complementary goods and services? Or will they be adopted too quickly, locking in the industry to a technology before the needs of most users have been clarified and addressed by product designers? Will the low (social) cost alternative always be chosen as a standard when tangible or intangible assets required for the production of vital technical components remain under exclusive ownership? Can de facto standardization on proprietary formulations
retard innovation? The apparent costs stemming from biases inherent in a dominant firm's decisions must be traded-off against apparent gains from the effects of greater firm size when it comes to coordinating and stabilizing product characteristics and production process specifications. Indeed, the issues surrounding the role of dominant firms are akin to those raised by consideration of the proper limits of natural monopolies; there is some societal gain when one firm supplies an entire system, or operates a network, because it internalizes the problems created by permitting incompatibilities (as AT&T sought to do prior to divestiture, and as IBM did in developing its System 360). In addition to the regulatory problems that natural monopolies pose, the trouble is that toleration of some level of transient incompatibilities may be essential for major innovations and for steady progress in the elaboration of complex technologies.

4. STANDARDS-WRITING ORGANIZATIONS AND REGULATORY CONTEXT

One indication of the perceived private and social gains from standardization is the increasing effort – much of which centers in information technology industries – to improve the performance of existing standards-setting bodies and, where that appears infeasible, to form new organizations. Whether created by the government as a public agency, or developed through private initiatives, these organizations appear to be growing more numerous and influential as information technologies play an increasingly important role in advanced economies (see, e.g., Cargill (1989)). More than 400 organizations have been estimated to be at work in the U.S., developing, revising, and reviewing standards (U.S. Federal Trade Commission (1983)).

Voluntary standards-setting organizations exist in many different product areas, each with its own formal and often complex procedures for developing standard specifications or reviewing proposed industry-wide standards (Hemenway (1975)). These groups differ in makeup with respect to the mix of the interests represented by the participants – and thus, in the primary motives of those who take part in their deliberations. For example, vendor firms may belong to some industry standard groups, whereas in others the actors are exclusively (professional) individuals (e.g., the members of the Institute of Electrical and Electronics Engineers).

In the U.S. the American National Standards Institute (ANSI), itself not a standards developing organization, develops and maintains the voluntary, consensus-based process to which all accredited standards-writing organizations must conform. Presently about 250 groups and 1000 companies belong to ANSI, and, in areas as technically distant as boilers are from process quality control, essentially the same consensus process is used to generate so-called American National Standards. ANSI also serves as U.S. representative to the International Standards Organization (ISO), the leading world-wide standards body. One should distinguish between national organizations and international bodies – such as the International Telegraph and Telephone Consultative Committee (CCITT), which is formed from the national members of the International Telecommunications Union (ITU) – because conflicts in the first arena typically involve disputes between firms with conflicting economic interests, whereas positions taken in the international standards arena may often be tied to larger national political and economic policies. Although personnel from corporations can act as national representatives in the deliberations of international standards bodies, government sanctioned policies such as the attempt to use stan-
ECONOMICS OF COMPATIBILITY STANDARDS

standards as non-tariff instruments to shape international trade may exercise an over­riding influence. CCITT develops recommendations for ITU, but the latter is a U. N. charter organization, in whose meetings it is government representatives who vote.

Within ANSI, the primary U. S. standards developing organizations for information processing and telecommunications, respectively, are called the X3 Committee, and T1, the Accredited Standards Committee for Telecommunications, which was formed after the breakup of AT & T. There are some 30 technical committees within the X3 organization, with specific fields such as digital and magnetic tape (X3B5), computer graphics (X3H4), codes and characters (X3L2), etc. Each technical committee comprises from 50 to 200 participants, actively engaged in the standards process and organized in sub-committees with narrowly specified tasks.

The influence of these organizations over the design, the timing, and the likelihood of adoption of specifications for particular standards varies widely. It depends on many of the features mentioned above, including (a) the degree to which the technology is already developed and its characteristics and cost structure are, therefore, widely known within the industry, (b) the degree to which investments in particular standards have already occurred, and (c) the degree to which relevant decision makers influence the process, as well as the internal political dynamics of the group (Hemenway (1975), Farrell and Saloner (1988), Weiss and Sirbu (1990)). Much of the economics literature about standards-writing groups endeavored to clarify the most important dimensions and the basic factors at work in the development of 'anticipatory standards', which only comparatively recently (during the 1980s) became a focus of the activities of information standards-writing organizations (see Weiss (1989)).

4.1. Voluntary standards-writing committees in practice

The development of an anticipatory standard can be understood as an exercise in collective engineering research and consensus-based product design. How do the technical committees of voluntary standards organizations in practice develop anticipatory standards – i.e., standards that are created ahead of a product? Usually the process begins with some reference model, such as ISDN (Integrated Services Digital Network) or OSI (Open Standards Interface), which guides the technical context in which the specific standards are developed. Then the formal process begins. This involves developing the new item, identifying all its services, and lastly, putting all relevant code down into a standard protocol. A less formal process follows in which a reference group, such as a user-group, considers the mass of options written in the standards and distill from it those subsets of the standard that are especially relevant to their group's members. After this, firms begin to develop both products and tests for conformity of products to the standard. Once a common language for testing is developed, the certification of products can proceed.

There are two fundamental concerns with the operation of this process (see Sirbu in David and Greenstein (1989)). First, standards-setting committees tend to write in old ideas and not new ones. It is difficult to develop within the consensus format any ideas that everyone does not already know about, at least partially. This biases technical development away from imaginative solutions. Second, decisions often are so technically complicated that only those who livelihoods depend on it can keep the complexity straight (Witten (1983)). Vendors know more about the technologies, but the debates often bog down in arcane technical issues that are inaccessible to many others, including some representatives of the user community.
Many changes are occurring in standards-setting bodies today to accommodate the increasing need for anticipatory standards (see below). For example, CCITT originally was an informal group for developing telecommunications standards. It has now grown into a legislative body, with funding from the UN, and its standards are enforced—by governments—to different degrees in different countries. CCITT procedures also have changed recently: (1) standards are now approved after a 70% consensus, rather than the 100% consensus that was required previously, and, (2) the time needed for approval has been shortened from four years to one year. Both changes can be expected to make CCITT standards-setting more responsive to current technological and commercial developments, but, since the resulting standards carry force in some national markets, there is potential danger as well (see Maxwell in David and Greenstein (1989)).

U.S. standards-writing organizations differ widely in the structure of decision-making and the delegation of authority (Cropper (1980), U.S. Federal Trade Commission (1983)). So, economists naturally have tried to identify which institutional features essentially shape committee process and which tend to be of secondary importance. In their case study of the X.25 standard, Sirbu and Zwimpfer (1985) found that standards are more likely, (1) when interfaces correspond to “natural” divisions in the responsibility for the supply of complementary equipment, (2) when there is practical experience with a technology and development was imminent, (3) when sponsoring firms built coalitions in favor of standards prior to joint proposals, and (4) when decision-makers with administrative responsibilities in the affected companies, rather than those who solely possess technical expertise, are involved in having the standard adopted. In their analysis of seven cases of standardization, Besen and Johnson (1986) identified several conditions that facilitated agreements on standards, such as are reached within the voluntary organizations’ committee process. They stress that: (1) all major industry parties must be willing to participate in standardization processes; (2) any industry group has to overcome potential antitrust objections, particularly to establish arrangements for considering dissenting views; (3) a group must find a way to narrow the range of choices, so interested parties can more easily arrive at consensus; (4) groups must develop objective technical means for considering alternatives, which facilitate agreements by eliminating subjective disputes on as many issues as possible; (5) liberal licensing arrangements and nominal royalty fees may reduce the costs to firms whose proposals were not accepted, and hence, make them more amenable to consensus.

More research is needed to draw connections between how standards committees operate in practice and what biases a consensus rule might impose on the development of standards in markets where these organizations are the main source of coordination efforts. This involves a deeper analysis of who it is that sets the objectives in standards-writing committees, how these objectives are modified, and how participants translate goals into actual designs. Research should address how behavior in committees that are developing standards in anticipation of products differs from that in committees that are developing standards in markets where de facto standards are already established. Empirical studies thus far have not managed to correlate institutional arrangements with committee performance in updating and revising standards, or identify differences between committee performance and corporate sponsorship in this regard. Such inquiries could help to identify what type of organization is best capable of producing anticipatory standards for rapidly changing technologies—as
opposed to efforts aimed at codifying and regularizing existing conventional design formulations and procedures.

4.2. Strategic behavior within committees

Voluntary standards-writing organizations are of analytic interest because they widen the number of channels through which specifications for standards can develop and resolve anticipated or extant technical conflicts between alternative designs. As a result of these channels, firms have available a potentially large number of strategic options and opportunities for influencing the eventual standard’s specification or the timing of its final announcement by taking actions in phases before the final vote for adoption by a standards-writing group.

Strategies available to firms participating in administrative structures can be quite complex to describe and document, since actions are taken in anticipation of expected market reactions to committee-generated standards. Likely opposing firm strategies in market committee processes may vary, depending on the technical and structural situation facing the firms. That is, strategies are likely to differ depending on the degree to which the technology is developed and its characteristics are known, the degree to which investments in particular standards have already occurred, and the degree to which relevant decision makers can influence the institutions in the process, as well as the internal political dynamics of the group.

Because of this complexity, there has been little theoretical examination of the role of standards committees in standardization processes. A notable exception is Farrell and Saloner’s (1988) analysis of committees as a mechanism to achieve coordination when participants differ sharply as to which of two mutually incompatible standards they support, though both are better off if they agree. This structure for the problem casts the players in a “battle of the sexes”, to use the game-theoretic term. Farrell and Saloner compare the coordination performance of a simple model of “market leadership” with their battle of the sexes version of talk within a committee, and also examine a hybrid of the two processes.

They find that the committee takes a larger number of “periods” (negotiation rounds) to reach a consensus than does the market, but tends to coordinate more often (and hence, do better, if no value is placed on speed). They also find that the hybrid of the two systems performs better than the pure committee process, essentially because the development of alternative paths for coordination in any time period increases the probability of successful agreement. Though the results are somewhat conditioned on the structure of the game (e.g. committee participants cannot draw up “compromise standards” and the market game has to be somewhat stylized), this line of analysis exposes some of essential mechanisms at work in the two types of coordination methods, and suggests possible policy recommendations about when markets are best left alone and when committees may improve on market performance. Further work along this line for different kind of situations and committee processes seems warranted, but some new approaches also seem called for. One cannot evaluate the comparative performance of the voluntary committee process for “anticipatory” standards within this framework, since there is no market game to be played before products have been designed; collaborative R & D among firms seems the more obvious alternative institutional arrangement. The empirical and policy relevance of the analysis would be enhanced also by establishing “natural” units for the lengths of committee negotiation rounds and bouts of market competition. If
these are substantially different, the meaning of Farrell and Saloner's (1989) results on the comparative speeds of committees and markets remains ambiguous.

Although the complexity of both the institutional arrangements and the technologies makes it difficult to carry out empirical studies of standards committees' behavior, some very interesting work has been done. In their analysis of the development of standards for local area networks, Sirbu and Hughes (1986) argue that the development of LAN standards within committees should be understood as more than just a resolution of conflicts between opposing parties with different economic interests. They argue that anticipatory standards-writing interacts closely with the innovation process, resulting in frequent disagreements between participants about the most important technological and marketing problems. Committee processes should be understood as forums for educating the participants to a common perception of the problems to be solved. Because of this technical uncertainty, firms that invested resources in mastering the process influenced it in directions favorable to their interests. Mastering the technical aspects was especially significant when standards evolved, and the relevant committees were obliged to consider additions, variations, and new options.

In order to shed light on the decisions of firms within committees, Weiss and Sirbu (1990) collected statistical data on 11 cases where a committee decided between two contending standards. They found significant differences in the coalitions of firms that favored and opposed what eventually became the adopted standard. The winning choice had more supporting technical documentation behind it (measured by the number of technical papers in favor), but tended also to be advocated by committee members from firms that were known to weigh market factors more heavily than technical considerations. Winning standards also tended to be ones that attracted support among users of compatible products, even though the vendors of such products remained indifferent. Proponents of both the non-adopted and adopted standard retained equally firm beliefs in the overall technical superiority of their respective candidates, even after the decision. This finding is consistent with Weiss and Sirbu's view that standards-writing committees have become an integral part of the product development process for companies that are oriented to markets in which network externalities are significant.

The respective roles of vendors and future users of standards require further delineation. A complete view must incorporate the observation that many standards provide technical solutions prior to precise identification of customers' problems, but this can undercut a vendor's ability to design proprietary solutions. It is unclear whether or not a situation in which no firm "owns a standard" leads to more significant cost reduction, or less product differentiation, than emerges from the rivalry among proprietary formulations of the product's technical features. This should depend on how the process mediates between vendors (who tend to know more about the technical capabilities and costs of a solution though not precisely the needs of users), and future users (who have a better feel for what is useful but may not understand the relevant costs). Hence, it should depend on the extent to which the resulting standard is "commodity-based" for general purpose use and the extent to which it reflects users' "wish-lists" (see Maxwell in David and Greenstein (1989)). A full view of the process must also integrate the follow-up roles of user groups, in drawing up "profiles" of standards and developing enforceable conformity tests.

In general, future work should attempt to assess the costs borne by private companies, and the incentives that appear to justify the resource expenditure entailed in
having personnel participating regularly in standards-writing groups. Beyond the
direct costs in personnel time, there is the resource commitment (including R & D)
that firms make to support such participation, and acquire influence over the out­
comes. How extensive such commitments are will depend on what firms hope to learn
and accomplish through participation in such groups, and what other options are
available to them - to divulge technical information unilaterally or hold it back, to
agree or disagree quickly with a proposal, or to stall, seek compromise or go to the
market. Moreover, the literature on coordination could be linked better to the
micro-institutional arrangements of the voluntary standards organizations - i.e., what
features of standards-writing committees help solve some of the coordination pro­
blems typically associated with market processes? Will more information in a
“cooperative” setting solve conflicts when preferences diverge, or will it exacerbate
problems because of incentives to exaggerate one’s preferences? And how should the
analysis of firm strategies be modified, if at all, when considering the standards-writ­
ing forum of the international organizations?

5. GOVERNMENT INTERVENTION

Government regulatory bodies may have an interest in standards-setting for many
reasons, either because some government agency (a) holds authority to regulate the
industry’s firms (e.g., as in the case of the FCC; see also National Bureau of Standards
(1977)), (b) perceives that the result of standardization activity affects important
national goals, such as protecting domestic employment or maintaining defense
capabilities, (c) concludes that voluntary industry-wide standardization activities
have had an effect that is improperly stifling of market competition (Rosenberg
(1976), U. S. Federal Trade Commission (1978, 1983)). In addition, much of the
research on network externalities and coordination problems has suggested that
intervention by a central government can solve one or another externality problem.
Arguments for central government intervention are also encouraged by the observa­
tion that adoption of a good with a network externality confers a public good on all
subsequent adopters (Dybvig and Spatt (1983)).

Such activity is important to the extent that it may result in mandatory adoption
of particular standards, and thus influence the particular specifications of a standard,
its subsequent modifications, and the timing of its adoption. Of course, settings will
differ in the probability of regulatory intervention to resolve standardization pro­
blems and the principles which guide the type of action government organizations
take, as well as how these actions are tailored to the circumstances of each situation
(Berg (1987), Besen and Johnson (1986)). Designing efficacious actions and appro­
priate guiding principles for every situation opens a large research agenda.

5.1. “Public goods” grounds for intervention

One analysis of government intervention emphasizes the “public good” aspects of
standardization. Many firms and users benefit from standards, but it is not im­
mediately apparent that any individual vendor has sufficient private incentive to
account for the benefits accruing to others when taking actions leading toward a
standard (Berg (1987, 1989a, 1989b), Lecraw (1984)). This observation has particular
force with regard to conformity testing (Tassey (1982)), where information regarding
a particular vendor’s product would benefit all buyers were it disclosed. According to
this argument, in the absence of cooperation by all buyers, one can expect a socially insufficient amount of investment in standardization activity and there is a justification for government to step in and provide support for it.

Despite the conventional force of this analysis, it has not been confronted with much of the evidence concerning the extent of voluntary activity. It would be interesting systematically to reexamine this argument's relevance in light of the efforts carried on by the myriad of voluntary organizations in this country (listed in U. S. Federal Trade Commission (1983)), and by close study of cases of apparently rapid "spontaneous" cooperation, such as the formulation and adoption of the Universal Product Code (bar code standard) among retailers (Keehn (1987)). To cite a countervailing finding from one such study, Berg (1989b), concludes that the adoption of technical standards by manufacturers of radio and television proceeded in seeming ignorance of important externalities, and did not take into account the likelihood of premature technological obsolescence, or the quantity of complementary broadcasts available to owners of receivers.

Further, and more difficult issues arise in a dynamic setting. David (1986a, 1987c) developed a conceptual framework for examining government intervention in de facto standardization processes when there are significant dynamic increasing returns. He highlighted three policy dilemmas regarding intervention when anticipatory standards are being developed: (1) there may only be narrow windows of time in which policy intervention can effectively influence the process, before the market "locks in" a standard (or the private and social costs of switching standards become prohibitively large); (2) at the time when government agencies can have the most influence, they also have the least amount of information about what action would be most appropriate - i.e., they are "blind giants"; (3) encouraging the development of gateways as a means to minimize the losses due to systems becoming orphans risks reducing the returns to system manufacturers from investing more in developing the capabilities of their proprietary systems. It may also induce excessive reliance on costly gateway devices among users, as Farrell and Saloner (1989) have shown in a formal treatment of the problem.

Because the issues are complex in practice, the role of regulation in standards-setting issues may present vexing problems to government agencies. Besen and Johnson (1986) examined several instances of FCC regulatory intervention in the development of standards in different markets (also see Sterling (1982), Berg (1984)). They focused on the question: when should a government agency choose to ratify market standards, try to influence the standards-setting process, or mandate a standard? Besen and Johnson stress the importance of the character of interactions between government regulators and industry participants, arguing that: (1) where the relevant technical deliberations already have taken place, a governmental agency should depend more heavily on industry evaluations, rather than conducting its own technical evaluations; (2) government should not mandate standards if these are likely soon to require revision - as tends to be the case where technologies are undergoing rapid change and an engineering consensus evidently is difficult to establish; (3) performance-oriented specifications are generally preferable to design specifications, especially in developing anticipatory standards; (4) symptoms of ineffective or premature actions should not be ignored - including negative industry reactions and continuing attempts to break from mandated standards; (5) sparse response to a proposal may indicate premature action.

While the arguments made in favor of government intervention on the grounds of
efficiency gains are well known, less frequently cited are some of the potential losses imposed by the regulatory process. First, some groups systematically acquire more influence than others because they are unequally represented when the issues are arcane. Even though government would like to require that all relevant parties be represented, not all parties can be identified. Second, the political process will exaggerate losses, especially when the losses or gains are in the future. It will also tend to protect old standards and accentuate identifiable "vested" interests. Thus, a standard may be politically easier to write when an activity is new, yet will be more difficult to write because it will not be clear what it should specify. Noll (in David and Greenstein (1989)) provides three examples illustrative of these principles.

5.2. Other government policy areas
An important topic that, surprisingly, has been overlooked by legal scholars and economic analysts alike concerns the influence on compatibility standards of the law protecting intellectual works. How ownership rights involving intellectual property are defined and delimited (and how such statutes are interpreted and enforced by the courts) partly determines the degree to which economic agents can influence the relative costs and benefits of particular choices of other economic agents. For example, the relative power of patent and copyright law to protect the ownership rights of rents created by technological innovation and legal norms for appropriate market conduct by owners of patents affect the supply of specifications which are potential standards. They also affect the extent to which existing specifications are imitated, adopted, and become de facto standards. This issue has become increasingly important in recent court cases where firms attempt to use the law to protect proprietary technologies embedded in standards. Farrell (1989) has recently addressed the general role of intellectual property law in market mediated standardization processes, noting that "weak" protection may encourage entry of imitators, creating price competition that favors de facto standardization; strong protection, on the other hand, may strengthen the hand of sponsors both in the market and in formal committee negotiations.

Several other contributions have examined how patent and copyright protection of interface standards and related technologies might be tailored to achieve optimal outcomes. One issue centers on whether proprietary rights should be granted for a particular standard interface or only for particular technologies. Menell (1987) argues that providing too much protection to inventors of software risks providing them control over complementary products. As has been noted above in section 3.5, some analytical support has been offered for this argument. While it is generally agreed that the use of interfaces which exclude rivals by raising their costs should not be allowed, the normative analysis be must done carefully when technical tie-ins are used to achieve price discrimination. If price discrimination were to be permitted, then it should be permissible to achieve discrimination by contract. Firms should not be induced - by the absence of contractual instruments - to undertake socially costly and otherwise pointless manipulations of technology for the purpose of achieving price discrimination (see Baxter in David and Greenstein (1989)).

In regard to the specific case of extending copyright protection for computer software, Farrell (1989) has noted that the legal principle of protecting not the "idea" but its specific mode of "expression" is poorly suited to encouragement of standardization of user interfaces: software producers are encouraged to differentiate the
outward features of programs so as to create a protected asset, and the creation of interoperable programs may be discouraged.

Another policy area that has been tentatively explored concerns the ways in which standards organizations can be captured by incumbent firms, and thus serve to facilitate anti-competitive behavior. Some observers have been led to propose, therefore, that standards committees write only performance standards and not design standards. While antitrust policy should be permissive regarding standardization activities that encourage entry and competition, the practical principles for regulating behavior in such organizations and balancing all relevant concerns are difficult to articulate (U. S. Federal Trade Commission (1978, 1983)). This is particularly so in network industries where cooperation can yield large efficiency gains (Carlton and Klamer (1983), Braunstein and White (1985)). For example, although some observers propose that standards committees only write performance standards (because entry barriers are lower and innovations are easier to adopt when a standard does not specify exactly how the product must work), performance standards will not suffice to achieve interoperability in many situations (Rosenberg (1976)).

Lastly, governments have used standards, like other policies affecting international commerce, as instruments to promote national objectives— even when these are in conflict with securing a Pareto-optimal outcome at the global level (see for example, Bar and Borrus (1987)). This is increasingly a concern as private industry groups in the U. S., such as the Corporation for Open Systems, find themselves negotiating with European and Japanese firms over standards that will have significant consequences for the entry conditions facing importing firms (see Metcalfe in David and Greenstein (1989)).

The most studied example of standardization with international repercussions concerns the attempts (and ultimately the failure) to achieve worldwide standardization in the production of television equipment (Crane (1979), Pelkman and Bueters (1986)). This failure can be traced to the contrasting objectives of national strategies, the private interests behind them, and the lack of incentives for decision-makers to account for user-community interests. Indeed, previous decisions regarding standardization in televisions derive from a national interest to use standardization decisions as a means to achieve external trade protection. One senses from these analyses that the impending attempts to standardize high-definition television production will be subject to very similar factors.

Considerable experience has now been achieved in the public administration of standards. Surveying the areas in which public agencies have set standards is one way we may cut across, and thus draw connections among, bodies of expertise that otherwise might be quite isolated. For example, the U. S. Federal government has been overseeing the operations of telecommunications companies for many years, but only recently has been concerned with the changing international competitive environment for telecommunications equipment. Technical standards in the purchase of computer equipment have been an ongoing concern of the General Services Administration, but only recently has the Federal government attempted to promulgate a standard for computer language (the Department of Defense’s sponsorship of the language ADA). A survey of this experience would also be useful for summarizing both prominent successes and failures of past efforts.

Research has made strides in understanding when an agency with regulatory power would do better to leave standardization to the market or to voluntary committee processes, but, a deeper analysis of the latter two is required as a basis for implementing some of these insights in concrete situations.
ECONOMICS OF COMPATIBILITY STANDARDS

6. CONCLUSION – RESEARCH ACHIEVEMENTS, NEEDS AND OPPORTUNITIES

Economists taking up the subject of technological standards, by and large, have proceeded by examining the circumstances surrounding specific technologies and industries in a case study mode, or by analyzing the implications of quite specific formal models. Yet, from the diversity of the particular cases and models upon which economic analysis has focused, several major unifying themes have emerged.

Market guidance of technical choices between competing technical designs, unlike the provision of market incentives for investment in technological innovation, was a matter that neoclassical analysis tended to view as essentially unproblematic. For some time, at least since the work of Hemenway (1975), an exception was recognized to arise from the “public goods” nature of an established technical standard, whether of the minimum quality, reference, or interface variety. If providing these was a costly activity and it was left solely to the workings of the competitive market, free-riding behavior on the part of profit-maximizing firms supposedly would tend to result in too much diversity and too little coordination. Here was ground for public intervention, if it could be presumed that the public authority was at least as technically well-informed as a private monopolist—who would be motivated to act in order to capture the efficiency gains from standardization.

A newer theme that runs through the recent literature concerns the multiplicity of plausible outcomes that exist where there are appreciable economic gains simply from achieving coordination among the actions of different agents. This is the fundamental situation that has been taken to characterize the problem of setting technical standards to insure compatibility or interoperability of system components. Correspondingly, it has been recognized that the attractions of sharing in “coordination benefits” may be a source of unexpected and sometimes perverse outcomes when standardization is achieved de facto, by market rivalry among variant formulations of a technology; or when standards-writing organizations are influenced by on-going market competition among products embodying alternative design concepts. This is the phenomenon often referred to as “bandwagon effects”, or “excess market momentum”. The latter have been found to be engendered in economic processes more generally by positive feedbacks (see David (1988), Arthur (1990)). Broadly speaking, “positive feedback” is a type of complementarity relationship that operates over time, a dynamic form of “increasing returns” in which the occurrence of one action (or event) raises the likelihood of that action being taken (or occurring) again. Underlying such positive feedbacks, which cause a particular standard to become more attractive to potential adopters as it gains adherents, there usually will be found one or another specific mechanism of self-reinforcement or collateral-reinforcement.

Thus, during the past decade economists’ research focus has tended to shift from the older question of whether or not markets will fail to provide timely technological standards in sufficient number. It has come to rest more centrally on a different set of fundamental questions: under conditions of positive feedback, will markets result in a choice of standards that have socially optimal characteristics? Are compatibility standards peculiarly subject to systemic market failures affecting their initial formulation, and their subsequent evolution?

The development of unsponsored standards subject to increasing returns to adoption received the earliest attention, and thus, developed the most coherent set of research results. It has been seen that that research tended to emphasize several concerns: first, market processes can lead to the establishment of one of several
possible de facto standards, if any at all. Hence, the timing and character of a particular standard can alter many outcomes observed in the market place. In addition, market processes do not automatically result in an optimal choice of standards. Insufficient or excessive standardization are both possible and, in some circumstances, either may occur whether or not there is full communication among the interested parties. Finally, when decisions to adopt (or adhere to) a standard are taken sequentially, later users have a tendency to emulate the decisions of the majority of earlier users in markets characterized by consumption externalities. Thus, market processes which develop standards over time may be unduly influenced by the decisions of early adopters who did not consider their effects on later adopters.

Treatment of sponsored standardization processes comprise a much more eclectic body of research. Two topics dominate much of the literature that can be organized on this heading: (1) How does competition ensue between sponsored incompatible systems embodying different standards? (2) What is the incentive of a system sponsor to make its system compatible with a rival’s incompatible system? The first research line tends to emphasize that sponsorship per se may change the dynamics for standards development (relative to an unsponsored process). But, it is found that sponsorship is not sufficient to solve many of the problems associated with unsponsored standardization processes. Indeed, sponsorship may introduce concerns regarding sub-optimal market performance resulting from strategic behavior. Consistent with the latter point, most of the literature addressing the second question emphasizes that different market participants generally do not have the same incentives to be compatible with one another, much less optimal incentives. Their incentives diverge for a variety of reasons that have been reviewed above.

Several other sub-topics within the category of sponsored standardization have received attention primarily because they touch upon longer-standing concerns of the field of industrial organization. These include analysis of the causal relations between market structure and standardization processes (in both causal directions) and analyses of the ability of dominant firms to manipulate standards to their private advantage. As in unsponsored processes, many of these analyses turn on novelties that arise out of dynamic increasing returns, particularly how the timing and character of standardization choices shape industry evolution. In addition, there has been greater focus on how the technical interdependence between sub-components of a system provides a dominant firm with strategic options for competitive behavior that it would not otherwise have. Here, however, the details of the models make a difference and there is no strong consensus in the conclusions reached by different studies.

Far less attention has been devoted by economists and political economists to examining the workings of standards-writing organizations, and the internal processes of technology selection that characterize inter-firm consortia formed for the express purpose of developing and sponsoring standards. Much more must be learned empirically about the actual extent of the resources being committed to “anticipatory” standards-writing projects, and the consequences of the particular administrative policies and procedures adopted by these organizations. Certainly this is in order before outside observers from the social sciences can confidently venture putting forward proposals to reform the process. What work has been done has tended to emphasize fundamental economic concerns about the efficiency of standards-writing organizations, both as administrative units, and as agents for developing socially efficient standards. Most of this has aimed to develop normative guidelines for standards-writing organizations. Some work has also analyzed general principles for
guiding government intervention in the standards-writing process, emphasizing new issues that arise in areas of legal protection of intellectual products.

A great many research strides have been noted in the foregoing pages, advances both in the theoretical understanding of the dynamics of standardization processes and the characterization of empirical patterns observable in the actual operation of market and non-market processes. Yet the field remains young and our review has identified some empirical and theoretical issues that appear to be particularly deserving of attention in the future. It must be said, further, that in the first wave of enthusiasm for this subject, the modelling exercises have run well ahead of the solidly established fact base. Systematic empirical (and specifically applied econometric) studies in this field are only now beginning to appear, and the primarily empirical thrust of the contributions assembled in this issue of *The Economics of Innovation and New Technologies* represents a welcome development.

**Selected References**


Besen, Stanley M. and Leland L. Johnson (1986), *Compatibility Standards, Competition, and Innovation in
the Broadcasting Industry, Santa Monica, California: The RAND Corporation, November.


Bresnahan, Timothy and Amit Chopra (1990), 


Cowan, Robin (1988a), Narrow Windows, Blind Giants and Angry Orphans: The Dynamics of Systems


Economides, Nicholas (1989a), Desirability of Compatibility in the Absence of Network Externalities, American Economic Review, 79(5, December), 1165-1181.


Farrell, Joseph and Garth Saloner (1989), Converters, Compatibility and the Control of Interfaces. Working Paper Number 89-130, Department of Economics, University of California, Berkeley, December.


Farrell, Joseph and Garth Saloner (1985b), Standardization, Compatibility and Innovation. RAND Journal of Economics, 16(1, Spring), 70-83.


ECONOMICS OF COMPATIBILITY STANDARDS

Melnitsky, Benjamin (1953), Profiting from Industrial Standardization. New York: Conover Mast Publications.
Postrel, Steven (1985), Bandwagons and the Coordination of Standardized Behavior. UCLA Business School, October (mimeo).
Puffert, Douglas J. (1988), Standardization of Gauge and Integration of Railway Networks in Britain and America. Department of Economics, Stanford University, Spring (mimeo).
Puffert, Douglas J. (1987), Spacial Network Externalities - A Model with Application to the Historical Standardization of Railway Gauges. Social Science History Workshop paper, Department of Econom-
ics, Stanford University, Fall (mimeo).
Putnam, Hayes and Bartlett (1982), The Impact of Private Voluntary Standards on Industrial Innovation, Volumes 1 and 2, November.
Saloner, Garth (1990), The Economics of Computer Interface Standardization: the Case of UNIX. Economics of Innovation and New Technology, 1(1/2).
Sirbu, Marvin and Kent Hughes (1986), Standardization of Local Area Networks. Department of Engineering and Public Policy, Carnegie Mellon University, April (mimeo).
Sirbu, Marvin and Steven Stewart (1986), Market Structure and the Emergence of Standards. Department of Engineering and Public Policy, Carnegie Mellon University, October (mimeo).
ECONOMICS OF COMPATIBILITY STANDARDS


