NETWORK INTEGRATION VERSUS NETWORK SEGMENTATION

Eli M. Noam

1. The Setting[1]

The Integrated Services Digital Network (ISDN) is a major development priority of most advanced telecommunications administrations and equipment manufacturers. It is one of those subjects which are at once obscure, complex, costly, and important. For a long time, there was no policy analysis of ISDN, only technological discussions. In most countries, decisions about the telecommunications infrastructure were made almost entirely outside the public view by engineering bureaucracies in government and manufacturing firms. Even where international negotiations forced these discussions into the open, such as in CCITT's Study Group XVIII, the record of several years of meetings shows little consideration of the cost or economic issues involved.[2] But as service integration — both narrowband and broadband — ceases to be simply an R&D project and becomes a major social investment, it must also meet non-technical tests. Therefore, ISDN's impact on industry structures and its relation to general telecommunications policy requires analysis.[3]

This essay argues that the economic case for the principle of service integration has not been adequately made, and that ISDN's significance is not merely as a technology upgrade; it may raise entry barriers against competition. There are merits to a segmented network environment too, and the move in the U.S. towards organizational differentiation in an open network environment must coexist with the equally important trend towards technical integration.

2. The Concept of ISDN

The term "ISDN" encompasses several sub-concepts and some confusion exists about its primary rationale. It is, first, a movement toward end-to-end *digitalization*. As such,

Integrated Broadband Networks: The Public Policy Issues / Martin C.J. Elton (Editor)

it continues a development of several decades, accelerated by the development of computers from analog to digital electronics. A second effect of ISDN is that it can upgrade user access to the telecommunication network to a *higher data transmission* rate. In past decades, the increased use of data, together with the greater information needs of users, has led to a vast increase in data communications traffic. Data communications rose almost forty-fold during the period from 1970 to 1985, while voice communications and data processing increased about five-fold.[4] This created an important need for communications links of a high capacity, greater than could be provided by narrowband analog networks. For voice traffic, too, digital transmission is more efficient when concentration and manipulation are required.

The third element of ISDN is *integration* which is more problematic. The ISDN concept aims to put together separate communications networks into one unified "super-pipe". From the technologist's perspective, this is a more elegant solution than duplication or multiplication of networks. As long as ISDN implies no more than technological improvement, it is hard to find fault with this development which is part of a larger technological trend. ISDN, however, is more than upgrading; it tends to be part of the business and political strategy of some telephone administrations to consolidate telecommunications in one standardized network. "Integration" is not simply an issue of technology, but one of business and regulation.

3. Economies of Scope

The classical arguments in favor of integration are the benefits of "economies of scale" and the joint production benefits of "economies of scope". The latter occur because the duplication of equipment and personnel is eliminated; similarly, less spare capacity is necessary, under almost all assumptions of traffic probabilities, to handle peak demand loads. If one denotes C as cost, and assumes two services, voice service V and data service D, economies of scope exist if the joint cost is less than the two separate costs, that is, if the following inequality holds:

$$C(V + D) < C(V) + C(D) \qquad (1)$$

That this relation exists is generally asserted by ISDN proponents as a matter of a priori reasoning. But the same logic —that substitute and complementary products are cheaper if jointly provided — applies similarly to numerous product pairs. In the extreme, if one wants to eliminate all duplication, the economy should consist of a single giant and fully integrated enterprise. But this hardly makes sense. The significance of economies of scope is not necessarily as clear-cut a matter as it may appear at first.

Furthermore, the question is not how best to structure a new network, but how to best restructure an existing network. In the short and medium term, virtually all capital in the networks V and D is sunk, whereas a modification for the integrated network $(V + D)$ requires new investments I, as well as the premature retirement R of some equipment used for the separate networks. Hence, the test for integration must be expanded to require

$$C(V + D) + C(I) + C(R) < C(A) + C(B). \quad (2)$$

In the short term, the adjustment costs $C(I)$ and $C(R)$ can be quite large and offset the pure economies of scope. In the longer term, their significance diminishes as equipment gets replaced naturally. But there are costs to gradual integration, too, because it requires parallel technologies of different generations to coexist for long periods of time.

Another problem with the argument of economies of scope is its implicit assumption of static cost functions. That is, costs are defined by the relations $C(V)$, $C(D)$, and $C(V{+}D)$ without a provision for change. However, analog voice and digital data networks are partial substitutes. Data can be transmitted over the voice network after passing through a modem that transforms it into an analog signal. Voice, similarly, can be digitized and be transmitted over a digital data network. Some information can also be transmitted either in voice or data form, depending on the costs involved. Hence, separate specialized networks, when under separate control, can provide a competitive environment that can lead to cost reductions and technical innovation. In other words, the cost curves in a non-integrated environment can be dynamic; what starts out as the cost relation $C(V)$ and $C(D)$ can, under the pressures of competition, become $C'(V)$ and $C'(D)$, with the assumption that

$$C'(V) < C(V) \text{ and } C'(D) < C(D). \quad (3)$$

It would then be more accurate to require the integrated network to meet the condition

$$C(V + D) < C'(V) + C'(D), \quad (4)$$

the existence of which is much less obvious than is relation (1).

Another problem with economies of scope is their lack of generality when the number of services is large. Suppose that instead of two networks V and D there exist n different types of data networks — providing, for example, different transmission rates, different error rates, etc. — denoted by $D(1),....D(n)$. Similarly, let there be a third type of telecommunications network, used for transmission of television programs, and denoted by TV. Then we have the separate cost functions

$$C(V), C(D(1)),.......C(D(n)), C(TV).$$

It is far from clear where the economies of scope exist. Even if we assume that each service has economies of scope with its immediate neighbors, this does not prove that a total integration across the entire spectrum of networks would produce economies of scope relative to stand-alone provision. To prove the economies of total integration, one may have to show economies of scope between every pair of services, for example, between voice and television. Furthermore, each service may have its own and different economies of scale and a bundling may lead to sub-optimal pairing.

Even where total integration may be cheaper than stand-alone services, the integration may not be stable — *sustainable* — if some of the services drop out of the integrated package. To understand this, it is necessary to distinguish between three cost concepts: first, the stand-alone cost of service of type *i*, denoted by $C(i)$; second, the cost of operating *i* as part of a service package, which is $C(I(i))$; third, the increment to total cost $IC(i)$ by adding *i* to the service package:

$$IC(i) = C(I(1),\ldots I(n)) - C(I(1)\ldots I(i-1), I(i+1),\ldots I(n)). \quad (5)$$

IC may be lower than the actual resource cost of integration, owing to positive externalities of *i* on the other services; for example, it may act as a backup. Suppose that the cost of providing i as a stand-alone service is 10, and, as an integrated service, because of technical problems of integration, the cost is 20. This would suggest that integration is uneconomical. However, the incremental total cost of the integrated service may be only 5, as a result of the positive externalities on the other services. Thus, integration would be economically efficient, but it would not be a stable solution. If each network could be operated independently, *i* would drop out of an integrated network, since this would reduce its cost from 20 to 10. Its positive externalities to the overall service would then be lost to the network. Thus, an integrated system, if it were voluntary, would be unstable or *unsustainable*.

The *sustainability* argument for integration, presented here in very simple terms, is based on a more complete general theory of industrial organization.[5] The argument for *non-sustainability* depends on certain underlying cost function characteristics which may or may not exist.

4. Economies of Scale

We now widen the discussion to economies of scale, the intellectual ancestor of economies of scope. This concept, probably the single favorite idea in the intellectual armory of telecommunications administrations, states that the cost of providing for additional users continuously decreases, and that it is thus cheaper to service all users with one large entity. Whether such economies continue to be obtained beyond a certain size has been questioned on empirical grounds.[6] It can also be argued that the inefficiency of monopoly more than offsets its economies of scale. In the American context the traditional system of rate-of-return regulation created built-in incentives towards over-capitalization, known as the *Averch-Johnson effect*. As discussed earlier, competition can move cost curves downward which may result in greater efficiencies than result from movement along a downward sloping static cost function, which is what economies of scale are about.

In the context of integrated services, economies of scale are used to argue that the integration of services creates cost efficiencies by bringing various new communications services to the population at large. Whereas it would be uneconomic for a residential or small business user to have digital data links, it would become affordable in an integrated universal network. By making such services prevalent, their cost per user drops, which is also of benefit to the earlier large users.

If strong economies of scale exist in ISDN, one policy implication could be to subsidize the early stages of its growth. Even if costs cannot be recovered immediately they will eventually come down while benefits rise. Early users may deserve a subsidy because their participation lowers the cost for previous and subsequent users; without a subsidy, they may never sign up and start the chain of economies of scale down the cost curve. This argument is plausible, though it does not prove a case for integration or for governmental control. But it is true for almost all start-up operations in any line of business, whether public or private. To justify subsidies in the ISDN context, one needs to have information about the size of the economies of scale.

The start-up subsidy necessary for the service becomes an argument for protection against competition and for maintenance of monopoly. Indeed, a highly perverse incentive is built in, because the greater the required subsidy, the greater the political support will be for a monopoly status that protects the investment from competition. An expensive project helps raise not only the economic, but also the political, barriers to entry. This encourages the introduction of successive large projects whose deficits justify barriers against cream skimming.

5. Scale and Scope

Economies of scale and scope are used as an argument for the expansion of network integration to all classes of participants in the public network. As with most of the pro-integration arguments, it is valid only under certain conditions. The theoretical problem can be demonstrated with the following simple model. Assume again two types of networks, voice (V) and data (D), and two types of users, business firms B and residential households R. Service can be provided separately, or jointly across services $(V + D)$ or across user classes $(B + R)$, or across both services and user classes. Figure 1 demonstrates schematically a situation in which economies of scale and scope exist, and yet the integration is uneconomical. Each number corresponds to the cost of providing a service of type i to user class of type j.

Figure 1.
Costs

	V	D	(V+D)
B	8	9	16
R	11	12	21
(B + R)	16	20	35

In Figure 1, economies of scale exist in the vertical columns, where joint provision to customers lowers costs for voice service *V* from 8 + 11 = 19 to 16 and for data service *D* from 9 + 12 = 21 to 20.

Similarly, economies of scope exist horizontally in the rows of Figure 1. They reduce cost for firms from 8 + 9 = 17 to 16, and for residential users from 11 + 12 = 23 to 21. Furthermore, full integration — both across services and customers — combines economies of scale and scope and lowers total costs to 35, where they would be 40 in total separation, and 36 and 37 in partial integrations. This would argue for total integration. The economies of integration (scale and scope) are greater than each of the sub-economies.

This analysis is based purely on cost considerations and is devoid of any discussion of revenues, benefits, and demand. If the price for each of the four sub-categories is arrived at by simply dividing total cost equally, the price for each sub-service would be 35:4 = 8.75. This means that firms would now be charged for the integrated voice service 2 x 8.75 = 17.5, rather than the 16 of partial integration, or the 17 of full separation. Thus, they would pay more than before.

It would be possible to structure a pricing scheme that would distribute the cost savings to all types of usage. This would mean price discrimination according to usage type or user category, or both. Yet such price discrimination would establish incentives for arbitrage. Under certain circumstances no price vector could exist that would not make it preferable for some service to be provided separately (*non-sustainability*).

A second problem exists if one considers that the different service types are of varying benefit to their users. For example, the utility of data services may actually be quite small to residential users. Let us assume for illustration that the benefits of the four categories are given by Figure 2. These benefits also establish the maximum willingness to pay.

Figure 2.
Benefits

	V	D	(V+D)
B	8	9	16
R	11	12	21
(B+R)	16	20	35

In this situation it will increase welfare (benefits minus costs) to integrate voice and data on for firms, since benefits are 19 against costs which are 16 (from Figure 1). But it would not cover cost to integrate the residential service, the cost of which would be 21, while maximum revenue would be only 11+ 6 = 17. Similarly, joint provision of voice services for firms and residential users would make sense, but joint provision of data services would not. Any inclusion of residential data service would have a deficit of 16 - 21 = -5. The problem could be solved through an outside subsidy, as by using the total cost gains of integration (8 + 9 + 11 + 12 - 35 = 5) to offset the deficit, or by charging some or all users and services above cost, up to the limit of users' willingness to pay. This means that three of the service types do not pay less due to integration; to the contrary, they are likely to be called upon to pay more than before, since earlier they paid less than their benefit (i.e., they had a consumer surplus). On the other hand, the fourth service is benefited. An otherwise lost benefit of 5 is gained, though at a cost of a subsidy of 5.

Another pricing policy would be to eschew a subsidy from one *user* type to another, and to limit it to a subsidy within a category. This could be done by requiring the residential user to obtain a package of the two services, or none at all. Maximum total payment would be, from Figure 2, 11 + 6 = 17. Cost of the package would be 21, from Figure 1. This would not be sufficient, and hence an additional subsidy of 4 would be necessary, either from outside governmental sources, or, more likely, from the other user category. This creates an incentive to go beyond partial integration to full integration, in order to find additional services or user surplus for purposes of subsidy.

But in terms of welfare, is this the optimizing policy? Suppose that instead of integrating all four services, one integrates only three and forgets about data service for residences. This would generate costs of 28 (for the three separate services), and benefits of 30, for a total net welfare of 30 - 28 = 2, whereas in the fully integrated system net benefits would be only 36 - 35 = 1. In other words, aggregate societal benefit may not be improved by a cross-subsidy to a user whose benefits from the service are small, even where one can show economies of both scale and scope. Thus, technical efficiencies by themselves do not prove the case for the societal benefits of integration in the absence of consideration of user benefits (i.e., of demand considerations). It is not enough to show that there are *some* benefits to residential users from data service, which is what ISDN advocates do with anecdotal evidence. Instead, it is necessary to illustrate the magnitude of these benefits relative to their costs.

This discussion can be extended beyond a 2 x 2 matrix to include additional types of telecommunications services and additional user categories. What is intended here is only a sketch of economic arguments that can be made more rigorously. Economic reasoning suggests that a trade-off exists between variety and specialization and that an optimal product diversity exists.[7] There can be too much standardization or inefficient standardization[8] which locks equipment manufacturing, systems development, and user applications into a pattern revolving around the one standard which may soon be technologically obsolete; however, for any one party to move to another standard may prove prohibitively expensive. Standardization implies an element of discontinuity and it can retard innovation. Such discontinuities exist in particular when, as in a network, positive externalities are generated by each participant. To leave a standardized integrated system is costly, because those departing lose the benefits conveyed by the other members. Only large parties can be expected to move on their own towards a new technical specification. However, because their departure imposes a cost — lost positive externalities — on the remaining adherents to the network, the decision to move to another standard may be blocked by administrative fiat as imposing a social cost.

Similarly, the early announcement of new standards, such as those for ISDN, is not necessarily the efficient information-enhancing act claimed by its proponents. Instead, it can be the act of a monopolist or of oligopolists aimed at preventing users from moving toward superior technology which they would otherwise choose, by signaling to them that they will lose the benefits of leaving the coalition served by the major standard that the monopolist controls.[9]

The investment in a major upgrading of the public network has the effect of raising barriers to entry. First, it increases the required initial investment which a potential rival needs to match the upgraded technical capabilities of the existing network. Second, where there is a trade-off between fixed costs and marginal costs, as there often is, the latter are lowered by the investment, making it more difficult for a rival to enter and match marginal cost pricing. Third, by raising the initial investment, one can stretch the range of economies of scale (declining average cost), and thus of "natural monopoly". The trough of a U-shaped cost curve is shifted to a higher level of production by an increase in fixed costs.

The theoretical analysis offered above demonstrates that creating an economic case for network integration, on purely analytical grounds, is not as easy as is claimed by ISDN proponents. However, the battles over ISDN have never been fought on the grounds of economics.

6. ISDN and Monopoly

The concept of ISDN, by itself, does not require monopoly. Strictly speaking, all it means is that the same communication link is able to provide a range of digital telecommunications services. The concept could be implemented with multiple and competing ISDNs or by private customized ISDNs. However, the usage has, at least within the various national telephone administrations known as the PTTs, implied a de facto exclusivity; the abolition of duplicative networks is stressed as a main goal of integration. Economies of scale and scope are the supporting arguments. With such justification, the idea of permitting rival networks seems self-defeating and deeply at odds with the motivations for ISDN. In a competitive environment, on the other hand, the concept of ISDN cannot be one of exclusivity, neither geographically nor functionally, but must involve the interconnectivity of multiple networks, including ISDNs and other types of networks.

ISDN has always meant different things to different people, from a simple and partial upgrading of digitalization to more ambitious undertakings including video transmission. But these different definitions revolve around technology. It is also possible to classify ISDN on the basis of its *purpose*. This would lead one to distinguish, for example, between an "upgrade" ISDN and "hegemonical" ISDN. The former is a step in the technical evolution of telecommunications. The latter is part of a more general effort to maintain monopoly control.

7. Network Fragmentation and the Open Network Architecture Concept

In the United States, the idea that all telecommunications should be provided by one organization was never fully accepted. Even in AT&T's heyday, it shared the field geographically with thousands of independent telephone companies (covering more than half the country and 20% of subscribers) and functionally with the domestic and international record carriers. Deregulation and divestiture accelerated the segmentation of networks. The term *segmentation* is not a negative characterization, and it must be discussed briefly. It means alternative or specialized networks, usually controlled and operated by several entities, and usually interconnecting with each other. First to emerge were private and then public alternative long distance carriers; subsequently, rival local transmission began to take place, known as *bypass service*. This was accelerated by the emergence of *shared tenant services*, which provide resale of local bypass service and also provide competition in local switching. These developments led to yet another and still more far-reaching approach, known as *Open Network Architecture* (ONA).

Open Network Architecture expanded the concepts of service alternatives and network fragmentation into the very core of the networks, and lowered barriers to entry for rival and varied communications services. ISDN, in contrast, raises entry barriers by providing a highly integrated network. ONA unbundles, while ISDN consolidates.

ONA is a framework in which network components are disaggregated in such a way as to permit open access to any of them. ONA permits the use by outside parties (users, enhanced service providers, resellers, and operators of other physical networks) of the building blocks of their choice. Where any of the blocks could be provided more cheaply or better by a supplier, other than a telephone company, it could be substituted and combined with blocks or equipment of the local exchange company. In other words, competition would be created for the various functions of the central exchange switch by unbundling its multiple functions. To make such a system work, service providers could conceivably collocate their own interconnecting equipment on the physical premises of the local exchange company.

The third party service providers are partly a form of value added network, with competing physical networks on the local level, and partly simple resellers. In all of these functions they compete head-on with the local exchange companies who act both as retailers and as wholesalers of these services.

A wide array of complex regulatory questions need to be resolved in the process of establishing such an open network.[10] In Europe a concept similar to ONA emerged under the name of *Open Network Provision* (ONP). It is a much more modest attempt to harmonize conditions of access for value added service providers from the European Community member states.

The ONA approach is not incompatible with ISDN in terms of technology; there could easily be an ISDN-ONA, and in the United States this is likely to happen. But in terms of underlying philosophy, ONA is diametrically different from the concept held by ISDN's major international proponents. The policy challenge is to maintain an open environment for multiple carriers and service providers, while permitting the upgrade of the network towards greater functionality.

8. Conclusion

Networks serve the exercise of speech and the free movement of information that is essential in a free society. These information flows require an increasingly sophisticated apparatus, run in most countries along a tightly centralized governmental model. This tension between the basically authoritarian industry structure and its democratic function worked reasonably well when technology and applications progressed slowly. In the past decade, however, major changes in technology, the importance of information as an input in the production process, and the general trend toward a services-based economy have vastly enlarged the significance and variety of telecommunications. This has led many to question whether any single organization, as effective as it may be, could or should be solely in charge of this key task in the economy.

Whether "segmentation" or "integration" are optimal solutions for networks in this environment is a matter that cannot be determined a priori. It involves a trade-off of four economic principles in two combinations: the efficiency of specialized production and of a competitive environment, versus the productivity contributions of economies of scale and scope and the reduction of uncertainty. One cannot generalize about what works best.

Domestically, these tendencies create the need to assure that non-integrated networks are not artificially hampered. Internationally, the divergent philosophies which underlie network policies in several countries make it difficult to reach stable international agreements on network integration.

Few of the questions raised in this essay are asked by the proponents of ISDN, who tend to view it as merely a technical upgrade, that is, another step in the improvement of the communications infrastructure. Ultimately, ISDN and the general issue of integration pose much more fundamental questions of control over the telecommunications network, the nervous system of an increasingly information-based society and economy.

Endnotes

1. This essay is based on the more extensive 1986 analysis, Eli Noam, "The Political Economy of ISDN", presented at the XIV Telecommunications Policy Research Conference, Airlie, Virginia, April, 1986. Distributed at that time as Working Paper #121, Center for Telecommunications and Information Studies, Columbia University.

2. Anthony Rutkowski, *Integrated Services Digital Networks*, Dedham, MA.: Artech House, 1985.

3. The need for such discussion was raised in Tony Newstead,"ISDN: A Solution in Search of a Problem?", *Telecommunications Policy*, March, 1986, pp. 2-4.

4. Department of Commerce, NTIA, "Primer on Integrated Services Digital Network (ISDN): Implications for Future Global Communications," NTIA Report, Washington, D.C., 1983-138.

5. W.J., Baumol, J.C. Panzar, and R.D. Willig, *Contestable Markets and the Theory of Industry Structure*, New York: Harcourt Brace Jovanovich, 1983.

6. For a review of the empirical literature, see J.R. Meyer, et al., *The Economics of Competition in the Telecommunications Industry*, Cambridge, Mass.: Oelgeschlager, Gunn & Hain, 1980.

7. K. Lancaster, "Socially Optimal Product Differentiation," *American Economic Review*, 1975, Vol. 64, pp. 567-85; D. Hemenway, *Industrywide Voluntary Product Standards*, Cambridge, Mass.:Ballinger, 1975; Spence, A.M., "Monopoly, Quality, and Regulation," *Bell Journal of Economics*, Autumn, 1975, pp. 417-429.

8. Joseph Farrell and Garth Saloner, "Standardization, Compatibility, and Innovations," Rand Journal of Economics, Spring, 1985, Vol. 16, No. 1, pp. 70-83; Joseph Farrell and Garth Saloner, "Standardization and Variety," *Economics Letter*, January, 1986, pp.71-74.

9. Joseph Farrell and Garth Saloner, "Installed Base and Compatibility: Innovation, Product Preannouncements and Predation," *MIT Working Papers*, No. 411, Feb.,1985.

10. Herbert Marks, 1986, "Comparable Efficient Interconnection: Equal Access is Better." Paper presented at GIA Roundtable; Eli Noam, FCC ONA Filing, submitted as CC Docket No. 88-2 Phase 1, 1988; Eli Noam, "Pluralism of Networks and Pluralism in Regulation: Open Network Interconnection in a Closed Federal System?," in *New Directions in Telecommunications Policy* , ed., Paula Newberg, Durham, N.C.: Duke University Press, 1989.