Peeling the "Layered Regulation" Onion

Glenn A. Woroch
Executive Director, Center for Research on Telecommunications Policy
Adjunct Professor of Economics
University of California, Berkeley

The Appeal of the Layered Approach

Many complex, high-tech products and services can be usefully viewed as a series of "layers." Like an onion, all the layers work together to produce the end result, each layer in physical or logical contact with its two neighbors. Engineers have proposed layered models for the architecture of telecommunications networks, the Internet, and computer systems. Recently, policy analysts and legal scholars have embraced this same engineering architecture as a template for reform of telecommunications regulation.

The so-called "layered regulation" of telecommunications—a framework that emerged out of debates over Internet policy—starts with functional distinctions of the various layers, from the physical infrastructure layer at the bottom of the hierarchy to application software and content near the top. The independence of each layer from the others is recognized as a powerful means to promote innovation. The central principle of layered regulation holds that regulators can reduce barriers to innovation by entrants and incumbents alike by ensuring their activities remain unencumbered by what occurs in other layers. If, instead, regulation was to "cross" two or more layers, innovation at one layer could be hindered by dependence on decisions made at another layer. Such dependence would occur, for instance, if voice service that rides on the Internet Protocol layer as a packet-switched application (i.e., voice-over-IP) was regulated in the same way as voice provided over the transport layer of the public switched network (PSTN).

An important corollary of layered regulation claims that, since upper layers depend on lower layers, the lowest layer—typically the physical network infrastructure—is the locus of the greatest monopoly power, and therefore, is where regulation should be concentrated. In practice, this translates into a policy that quarantines the "last mile" of the network from participation in complementary service markets.

The layered approach to regulation has several attractive features, beginning with its observation that current institutions are inconsistent with today's telecommunications technology. Presently, different regulations are applied to different networks—the switched wireline network, the cable TV video network, the mobile wireless network and so on. Each of these distinct service "silos" is further sub-divided by state-federal jurisdictions and by residential-business distinctions. Such Balkanization is irrelevant in a world of digital convergence in which voice, video, and data are digitized and transported over media of all kinds: copper pair, coaxial cable, optical fiber, the radio spectrum, even the electric power grid. Carving up the market into distinct services impedes the competition made possible by digital technology. Regulation applied to each service will inevitably follow a distinct evolutionary path, resulting in asymmetric treatment of potential competitors, further distorting the working of the competitive mechanism.

The layered approach is also correct to emphasize incentives for investment and innovation as the key criterion in the design of regulatory institutions. Technological advances are occurring at a torrid pace, even relative to the impressive historical record of the communications sector. The technological interdependencies that arise with these services pose the danger that sluggish advance by one component can retard progress for the entire system. Layered regulation seeks to achieve rapid technical progress in delivery of the final service by removing an important obstacle to innovation in each distinct layer.

A Better Way to Slice the Onion

Despite several compelling features, layered regulation fails to adhere to some basic principles of economic regulation that may, among other implications, defeat its pro-innovation goals. A minimum test for government intervention requires that the improvement registered over the unregulated outcome—derived from constraining the abuse of monopoly power, internalizing spillover effects, or pursuing social goals—is not overwhelmed by the costs of implementing the regulation. Additionally, the proposed policy should be better than then next best, feasible alternative. It is not enough merely to identify desirable properties that flow from layered regulation; it must prevail and pass the test of full comparative analysis.

One place to introduce economic principles into the design of layered regulation is to use supply and demand to delineate service components—whether they are arranged in vertical stacks, horizontal layers, or a more elaborate mosaic. Only by accident will engineering layers coincide exactly with the economic definition of service markets. Even in engineering terms, distinctions between piece-parts of a network can be far from clear. In addition, engineering and economic distinctions may lead to conflicting conclusions. As an example, whether some short-text messaging feature is embedded in the chip set of a mobile phone or located in the server software at the mobile switching center matters a great deal to engineers, but is irrelevant to the mobile user, the arbiter of economic service markets.

Furthermore, assigning products to layers is not always obvious. Take the case of “broadband over power line” (BPL), a technology that transmits data over the electric lines. Assuming that BPL becomes commercially viable, equating the electric power grid with cable and telephone networks, and thereby justifying symmetric regulatory treatment, could cause collateral harm to the efficient operation, maintenance and modernization of the power distribution network. The source of this harm is the dual purpose of this physical layer and the inability of regulatory policy to isolate clearly the telecommunications function.

Nor are lower layers in the “stack” necessarily prone to monopoly as suggested by proponents of layered regulation. More often, when barred from engaging in service differentiation, physical infrastructure slips into commodity status and intense competition soon follows. This occurred in the personal computer industry for the BIOS (basic input-output system) that resides near the bottom of the PC stack, located between the operating system and microprocessor layers. In sharp contrast to the layers directly above and below that are dominated by Microsoft and Intel, respectively, the market for BIOS firmware is highly competitive.

By isolating the physical infrastructure layer and imposing restrictive regulatory rules on its owner, opportunities for an important source of network innovation may be lost. Quarantined to the physical layer, and constrained in its profitability, the infrastructure provider foresees insufficient return to justify the enormous investment that often is necessary to retrofit an embedded network for a new technology. As a result, innovation that could otherwise percolate through the layers and deliver significant advances in end-user services is never realized.

The Other Side of Crossing Layers

Layers that can be delineated from an engineering perspective could nevertheless exhibit strong scope economies. It is likely that such economies will only be realized through vertical integration that minimizes the costs of completing transactions at arm's length. Integration can also eliminate the “double markups” that occur when two or more firms exercise their market power at individual layers. More important than these static economies, however, vertical integration can erect strong incentives for investment when firms can also realize returns in the service markets of adjacent layers. This can be particularly beneficial in the case of innovative activity where the surplus generated by cost reductions and service improvements can otherwise be difficult to appropriate from a single layer.

The possibility that a firm operating in several layers of a network will foreclose un-integrated rivals—especially innovative start-ups—is a serious concern, but one that is easily overestimated. Such firms can, and do, employ pricing, bundling, quality discrimination and interface control to frustrate their rivals. However, it is also well documented that firms possessing market power have strong incentives to facilitate competition in complementary products. Taking another example from the computer industry, Microsoft aggressively encourages entry into software and hardware markets that complement its famous operating system, despite its monopoly position and despite its predilection to exercise that market power. The profit motive also discourages a dominant firm from entering another layer when a more efficient, more innovative firm could supply that complement, provided the dominant firm is able to capture some portion of the rents generated by that complementer.

It is more than a little ironic that proponents of layered regulation see digital convergence as a reason to "de-laminate" the phone network,[1] and then to proceed to call for a quarantine of the physical infrastructure layer. Such a policy will inevitably impair the competition among platforms that is enabled by digital convergence. In contrast, when infrastructure owners are able to adopt cross-layer strategies to gain a competitive advantage over another platform offering, the consumer will be well served.

Summing Up

The layered framework is a useful schematic to organize complex networks like the Internet and telecommunications, but it is easy to be seduced by this orderly view. Superimposing this framework on regulatory institutions inevitably parts ways with the fundamental economics of these markets. While it holds great promise in promoting independent innovation, when we peel back the layers of the argument, the layered approach forgoes significant benefits that would only be realized by "violating" the layers.

Another drawback of imposing this engineering architecture on regulatory policy is that technology can change, and can do so quite quickly, while regulatory institutions are notoriously slow to react and to adjust. Incorporating the horizontal approach into institutions builds in rigidity that prevents regulation from adapting to the very innovative technologies that it seeks to promote. Who could confidently claim that technological advances will not occur in the future that would once again justify vertical silos?

Policy makers committed to facilitating innovation in modern digital networks have better alternatives at their disposal. These alternatives hold the promise of striking a balance between erecting profit incentives for inventing and deploying advanced technology while at the same time curbing monopoly power. For instance, the creation and dissemination of nonproprietary technical standards, especially open interface standards, along with obligations to interconnect and to maintain technical and commercial interoperability, would go a long way to promoting innovation while limiting the opportunity for abuse of monopoly power. Fortunately, institutions to support such arrangements would not have to be invented from scratch: the computer industry has ample experience implementing these concepts, and to a lesser extent, so too does the telecommunications industry.

---

[1] I borrow this term from Michael Katz, "Thoughts on the Implications of Technological Change for Telecommunications Policy," in Transition to an IP Environment, A Report of the 15th Annual Aspen Institute Conference on Telecommunications Policy, Aspen Institute: Washington, DC, 2001.