Wesleyan Economic Working Papers

http://repec.wesleyan.edu/
Nº: 2017-001

Infrastructure and General Purpose Technologies: A Technology Flow Framework

Christiaan Hogendorn and Brett Frischmann

January 2017

WESLEYAN UNIVERSITY

Department of Economics
Public Affairs Center
238 Church Street
Middletown, CT 06459-007

Tel: (860) 685-2340
Fax: (860) 685-2301
http://www.wesleyan.edu/econ

Infrastructure and General Purpose Technologies:
A Technology Flow Framework

Christiaan Hogendorn
Wesleyan University

Brett Frischmann
Cardozo Law School

January 18, 2017

Abstract: Economic growth models often refer to "general purpose technology" (GPT) and "infrastructure" as key to improving productivity. Some GPTs, like railroads and the Internet, fit common notions of infrastructure, while other likes the steam engine and the computer do not. Without a specific model of the special characteristics of infrastructures, important technology policy questions relating to openness are not addressed. Infrastructure is similar to other GPTs in its demand-side characteristics that enable a wide variety of productive activities (or uses) and generate substantial spillovers to the rest of the economy. On the supply side, infrastructure is quite different from the other GPTs. It is partially nonrival as opposed to fully nonrival, which may complicate appropriation problems and raise congestion issues. It has strong cost-side economies of scale giving less scope for using markets to provide and control it. And it exhibits tethering, meaning that different users must be physically or virtually connected for the infrastructure to function. We present a technology flow framework that clarifies these issues and provides a base for policy analysis and for defining empirical research questions.

Keywords: infrastructure, general purpose technologies, institutions, economic growth, spillovers, open systems

1. Introduction

While there is widespread agreement that "infrastructure," "general purpose technology" (GPT), and "open systems" are important to economic welfare, the connections between these concepts are often murky. We propose a conceptual framework that makes these terms clear and provides a consistent method to construct economic production functions and to motivate empirical research. The term ICT, for example, is widely used to group together computer hardware, software, the Internet, and telecommunications, but it also serves to avoid the complex interrelationships between these technologies. Our technology flow framework gives a rigorous method to choose which of these should be labeled as infrastructure and/or as a GPT.

The most important policy contribution of our framework is to clarify the "openness" of a technology or infrastructure facility. There has been a confusion of definitions in recent controversies surrounding network neutrality on the Internet, compatibility of different music and video players, and interconnection of telephone, electricity, and transportation networks. These debates frequently ask whether efficient use of infrastructure or technology can proceed from market forces alone, requires government intervention, or whether a third way involving commons management is possible. While our framework cannot decide a policy outcome, it does focus discussion on the precise elements of the technological system that require policy attention.

Most of the resources we think of as traditional "infrastructure," including electricity, railroads, roads, and the Internet, are also included in lists of what economists call general purpose technologies (GPTs). But economists typically count other technologies among GPTs – such as the steam engine and the computer – that are not usually labelled as "infrastructure." These labels matter. For example, the FCC used these ideas to justify its *2010 Open Internet Order* on network neutrality: "Like electricity and the computer, the Internet is a 'general purpose technology' that enables new methods of production that have a major impact on the entire economy. The Internet's founders intentionally built a network that is open, in the sense that it has no gatekeepers limiting innovation and communication through the network." (FCC 2010, pg. 5) This line was quoted and the argument was reconfirmed in the *2015 Open Internet Order* (FCC 2015).

This paper untangles the relationships between infrastructure and GPTs by combining two approaches from the literature on commons and commons management. From Charlotte Hess and Elinor Ostrom (2003) we use a setup where information goods (and we suggest other goods) are based on an *idea*, packaged as *artifacts*, and distributed through *facilities*. From Brett Frischmann (2012) we use a "demand-side theory of infrastructure" in which the uses – and thus the *demand-side* characteristics – of infrastructure, are its central defining features. Combining these ideas, we argue that most traditional infrastructures and all general purpose technologies have similar demand-side uses and effects on society, but that the scale and design of supply-side infrastructure facilities can inhibit open commons management.

In the next section, we describe the sometimes-confusing terminology and the concepts around infrastructure and GPTs. In Section 3 we build a technology flow framework for looking at the supply-demand relationships between technologies, and in Section 4 we give two important examples. In Section 5, we use the framework to specify the differences and similarities of infrastructure and GPTs, and in Section 6 we use the framework to address the openness question. In Section 7 we present some additional implications of using the technology flow framework, and we conclude in Section 8.

2. Infrastructure and General Purpose Technology

The facilities we call "infrastructure" are key to development and growth today. Trends such as globalization, integration of economies, and outsourcing are enabled by and interact with key infrastructure. From the outset, we need to confront a surprising difficulty: economics does not have a generally accepted definition of infrastructure, and there is not a substantial literature on infrastructure economics *per se* (Frischmann 2012). Instead, the importance of infrastructure and its policy issues are subsumed within development economics when they involve developing countries, regulatory economics when they involve regulated industries, macroeconomic growth when they affect the economy overall, and public economics when they are financed or franchised by the government.

There is, however, a recurring political and policy issue related to the methods by which infrastructure can be used. Here we label this the "openness" question, and it relates to how much control the owner of infrastructure can exercise over particular types of use of the infrastructure and, in some cases, over users themselves. Currently, the most active debate concerns local broadband Internet and goes by the name "network neutrality." But the same types of questions have occurred in railroads, telephone, and most other types of networked infrastructure (Hogendorn 2005).

Closely related to "infrastructure" is the term "general purpose technology" introduced by Bresnahan and Trajtenberg (1995). They argued that GPTs have three main features: pervasiveness, technological dynamism, and innovational complementarities, and they highlighted three examples, the steam engine, the electric motor, and the semiconductor. The Bresnahan and Trajtenberg paper sits on the cusp between microeconomics and macroeconomics, but the majority of follow-on research tilts to the macroeconomic side, particularly in the area of growth theory.

In their authoritative book on GPTs, Lipsey, Carlaw, and Bekar (2005, pg. 98, hereafter "LCB") present a specific definition of a GPT: "a single generic technology, recognizable as such over its whole lifetime, that initially has much scope for improvement and eventually comes to be widely used, to have many uses, and to have many spillover effects." The most far-reaching of these also have a transforming effect on society at large. They propose a list of these transforming GPTs (pg. 132, Table 5.1), which reveals that some are typically considered "infrastructure" while others are not:

Domestication of plants	Three-masted sailing ship	Motor vehicle
Domestication of animals	Printing	Airplane
Smelting of ore	Steam engine	Mass production,
Wheel	Factory system	continuous process, factory
Writing	Railway	Computer
Bronze	Iron steamship	Lean production
Iron	Internal combustion engine	Internet
Waterwheel	Electricity	Biotechnology
		Nanotechnologry

Table 1: List of "Transforming" GPTs from Lipsey, Carlaw, and Bekar (2005, pg. 132)

The concept of GPTs and lists such as the above have been criticized by Field (2008) for a lack of parallelism. Field argues that some commonly-cited GPTs are based on a broad definition – they produce an intermediate good that is used by many industries – while others are based on a narrow definition – they are directly used within many industries. Railways, for example, produce a very widely-used intermediate good, but there is essentially only one industry that employs the set of railway technologies. Thus the railway may be a GPT by the broad definition, but by the narrow definition it is a special-purpose technology that enables one particular mode of transport.

This need for parallel development is important in GPT models. In the list above, ships do little without ports, motor vehicles require highway systems, and airplanes require airports and air traffic control. Helpman and Trajtenberg (1998) model diffusion of GPTs by endogenizing the development of the components of GPTs in many sectors. They show there are multiple equilibria, with all, some, or none of the sectors coordinating around a new GPT. They do not, however, make any distinction as to whether the components are infrastructure or not.

The idea of parallel development of components is related to "poverty traps," which originates with Rosenstein-Rodan (1943). He said that the high fixed costs of modern technology require a "big push" by government to move multiple sectors of the economy to advanced technologies. Murphy, Shleifer, and Vishny (1989) modernized the idea and laid down two key conditions for a poverty trap: (i) investment increases the size of other firms' markets or increases the profitability of investment, and (ii) investment has negative NPV. This is consistent with the idea that infrastructure is complementary to many other sectors but has negative NPV when produced by a private firm due to its extensive spillovers (Frischmann and Hogendorn 2015).

Despite the importance of diffusion in the GPT literature and the presence of so many infrastructures among lists of GPTs, the specifics of infrastructure complements versus other types of complements has not been specified. There also has not been much discussion of the openness of GPTs to new uses – rather it is usually assumed. We believe this is primarily because most GPT research proceeds using macroeconomic growth methods, while infrastructure is treated as a microeconomic question for regulatory policy. We aim to build a framework that unites the two.

3. The Technology Flow Framework

We have seen that "infrastructure" is imprecisely defined and that GPT is precisely defined but subject to some confusion regarding categorization. We will use two tools to relate and refine these concepts: the economics of the commons and the demand side theory of infrastructure.

Hess and Ostrom (2003) treat information goods as a common pool resource. They discuss three important aspects of an information good: the artifact, the facility, and the idea (pg. 129). They define an *artifact* as "a discreet, observable, nameable representation of an idea," and give many examples including books, maps, and web pages. Artifacts are the physical (or virtual) resource flow units in an information commons. A *facility* "stores artifacts and makes them available" and is analogous to a commons' "resource system." Examples include libraries, archives, and online repositories. *Ideas* are "the creative vision, the intangible content, innovative information, and knowledge" contained in an artifact. They are another type of resource flow unit, but they are not tangible, even in a digital sense, and are not protected by copyright.

We believe that this trichotomy is useful beyond information goods for all resource systems including technology and infrastructure. This dovetails with Arthur (2012) and others who refer to the idea behind a technology as a "principle." LCB specifically call the idea behind a GPT a "general purpose principle" (GPP) and note that technologies come in "triplets:" product, process, and organization. This is a similar breakdown to Hess and Ostrom's, but we use the latter because it focuses more on the degree of rivalry of each element. To take one example, the system referred to as "electricity" involves the *idea* of electromagnetism, the *facility* of the power grid, and the *artifact* of electrons moving through the grid measured in kilowatt-hours.

Our other tool comes from Frischmann's (2012) book on infrastructure. He emphasizes the importance of a "demand-side theory of infrastructure" in addition to the more traditional focus on the supply-side. He argues that infrastructure is a partially nonrival input into a wide variety of productive activities that generate private, public, and nonmarket goods. The term "partially nonrival" is important – when demand is off-peak, output is nonrival, but on-peak demand can create congestion and thus rivalry.[1]

Frischmann says that infrastructure is a "shared means to many ends" (pg. 4). For example, a road can be used by cars, trucks and motorcycles, which we can think of as analogous to Internet packets. It is on the demand side that these different vehicles (and different origin-destination pairs, different cargoes, etc.) create diverse sources of value.

Our model of a technology uses the supply/demand distinction from Frischmann. Our main contention is that both the supply and the demand sides can be divided, separately, into the three Hess-Ostrom elements of idea, facility, and artifact. Supplying a technology requires a nonrival idea or principle, it requires some kind of enabling facility which is partially nonrival, and it then produces an artifact which is a rival, private good that we denote x_1. A demand-side application technology also requires a nonrival idea for a use or

1 Frischmann also discusses some "intellectual infrastructure" that is purely nonrival, but our model in this paper does not aim to incorporate those examples.

application, a partially nonrival facility to enable use, and an output which is a rival, private good denoted x_2. In the end, the two artifacts are combined into a useful, final output y, and this combination is itself an artifact.

The technology flow framework maps directly to the traditional production function $y=f(x_1,x_2)$, where the form of the function $f()$ is determined by the exact mechanism through which the combination of x_1 and x_2 takes place. That mechanism, in turn, is usually determined by the nature of the facilities involved, as we discuss in detail below. We illustrate the technology flow framework in Figure 1.

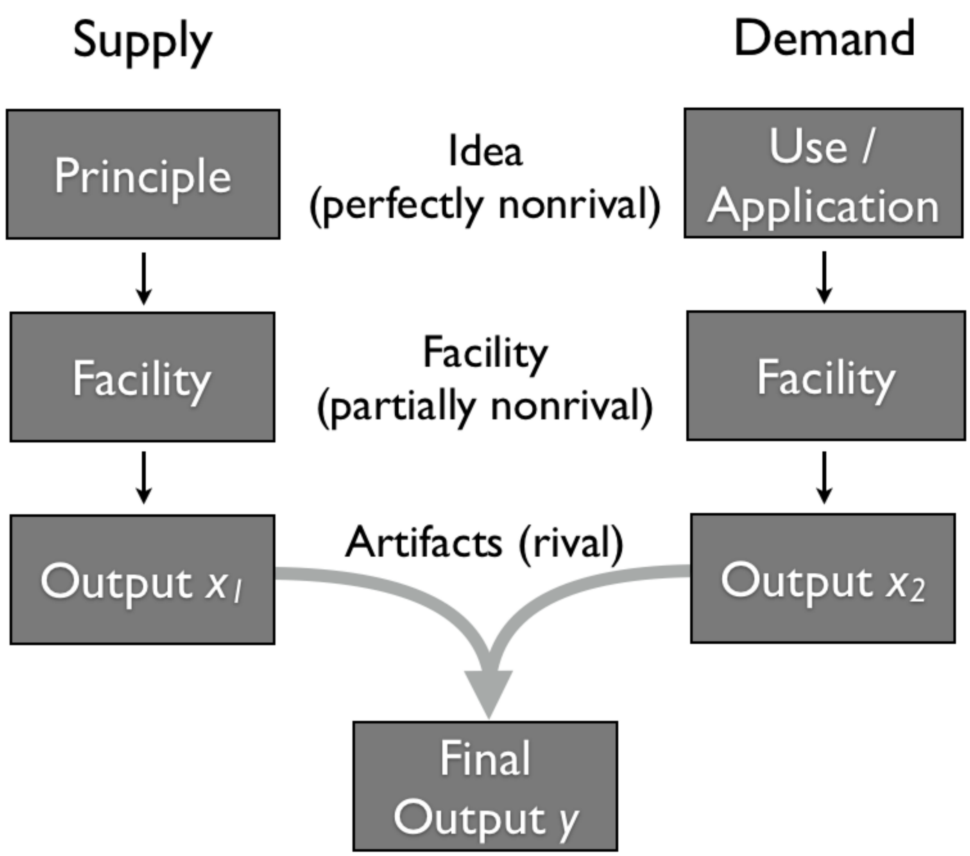

Figure 1: The Technology Flow Framework

It is easiest to understand this framework through examples, which we proceed to shortly. But let us note three important details before moving on.

First, we can use the technology flow framework to evaluate *any* technology, no matter whether anyone would call it general purpose or call it infrastructure or not. As LCB, Bresnahan and Trajtenberg, and others note, what makes a technology a GPT is its uses. If there are a variety of uses across a large range of the economy, that is what makes a technology general purpose. And as Frischmann notes, general-purpose infrastructure also has a great many uses. These authors also discuss the existence of special-purpose technologies. Thus the use of the term *general* refers to demand-side uses, and generalness is very important for a technology's role in the economy.[2] Nevertheless, as long as a technology has at least one use, it can fit in the framework.

Second, every technology is in fact an interdependent combination of other technologies, as many authors have made clear including Rosenberg (1978) and Arthur (2012). That means that one can nest the technology flow framework over and over, and in multiples, as in Figure 2. Every use of one technology can become a technology in itself that leads to other uses, and so on in long progressions. As with any model, the task of the researcher using the technology flow framework is to choose which technologies to highlight and which to abstract away. For example, display screens are essential to most digital technologies, but they can safely be left out of most models because they do not figure in technology policy questions.

Third, questions of *who* makes the final combination of x_1 and x_2 and *where* they do it touch on the theory of the firm and become very important to the openness question. In some cases the combination is made by the end-user, in other cases in the facility of the supply or demand technology, and in some cases the end-user must procure further services which may result in a multi-sided market situation.

2 LCB also note that "generalness" is measured on a continuum so that there are many "almost-GPTs" that are important technologies but not used so widely across the economy as to be important from a macroeconomic perspective.

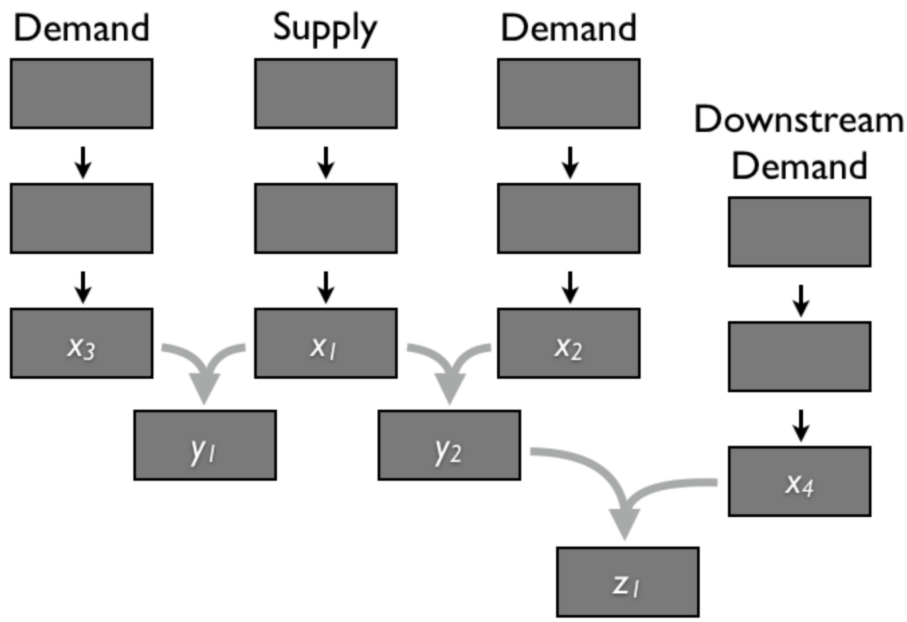

Figure 2: Extended technology flow relationships

4. The Computer and the Internet

Let us now turn to two specific examples. Rather like the steam engine and the railway, many lists of general purpose technologies include entries for both "the computer" and "the Internet." Both may pass muster as GPTs, but the argument for the Internet as "infrastructure" seems stronger than the computer. We think this is correct, and our framework shows why: while both are widely used on the demand side, the facilities to supply the Internet are much less rival than those that supply the computer.

Let us apply the framework to the traditional personal computer. The idea or principle behind it could be termed "electronic computing." It is the set of scientific principles and engineering know-how that allow a computer to be designed and factories to be built that can make computers.[3] The facility to make computers is a factory. More accurately it is a bundle of facilities that include not only factories but the design, marketing, service, and other functions that are usually bundled as a firm. Dell Computer is an example of one

3 As always, one can further subdivide, and say that computers are built on microprocessors, storage devices, and so forth. Older GPTs like electricity and even the wheel (on the mouse or disk drive) play a role in computers.

such facility organized as a firm. Ultimately the output created by Dell and other computer makers is an artifact, a single computer which is clearly a rival, private good.

An essential characteristic of computer production is that there is some rivalry. As a manufactured good, individual computers incorporate discrete chunks of materials, energy, labor time, and space. This means the facilities to create them use production functions with relatively high marginal costs. Computers also incorporate high fixed costs of development and support, so the final good embodies a share of nonrival inputs as well. We could thus say that the computer production function exhibits partial nonrivalry, but not to a great extent.[4]

For the most part, users do not want a computer as a final consumption good. Rather, there are many many different uses for the computer. One traditional use is the spreadsheet application, the original "killer app."[5] Again, this is based on a nonrival idea – the virtualization and automation of older paper methods of business analysis. The facilities to make a spreadsheet application generally involve a team of software developers and some source code – the Excel team from Microsoft is one example. The output of this team is a copy of Excel, another artifact. The two artifacts are then combined to create a working, installed spreadsheet application.[6] This example is illustrated in Figure 3.

4 An even more pure example of rivalrous production would be Adam Smith's pin factory, where the user of a pin would not need any design or support services at all.

5 Apparently the first reference to this term is Ramsdell (1980) regarding the VisiCalc spreadsheet.

6 The adjective *working* is important here. One could have a non-working copy of Excel even without a computer, just as one could have a non-working car without an internal combustion engine.

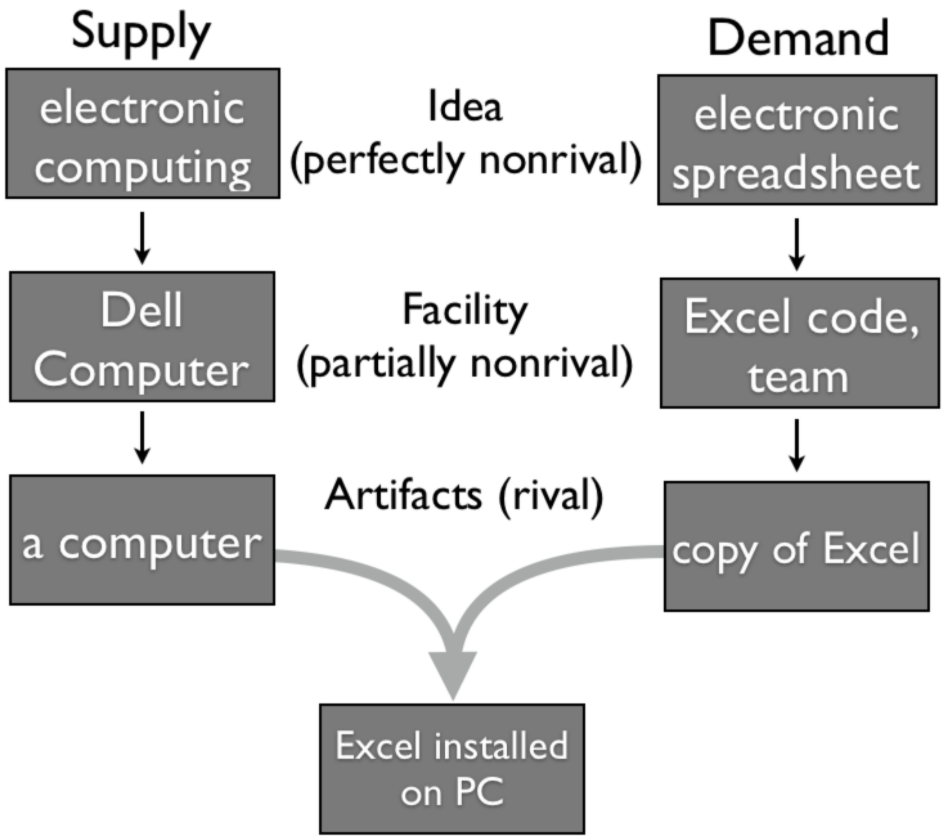

Figure 3: The computer and one traditional use, the spreadsheet

In the traditional desktop environment, the end user combines the computer artifact and Excel disk artifact at his or her own premise. Neither Dell nor Microsoft are involved in this final installation. This allow for some freedom to the user, who might, for example, operate the computer using Linux and install Excel into a virtual machine rather than the standard situation where Windows is the computer's operating system.

Now we turn to a technology closely related to the computer, the Internet. The idea or principle behind the Internet is packet-switched interconnection of computers. The facility is a group of interconnected networks including Internet backbones and local ISPs. The artifact is a packet of data sent over the network.

Internet packets are produced under partially nonrival conditions. There is very little marginal cost except under peak conditions, but the only way to provide these scale economies is for the user to maintain a direct, ongoing connection of the packets to the facility that transmits them. This is a general principle: the more nonrival is the production process, the more important is an ongoing connection to the facility.

From the demand side, there are again many uses of the Internet. The one most associated with the network neutrality debate is streaming video, which not only requires an ongoing connection to partially nonrival packet transport, but also an ongoing connection to the partially nonrival server capability of a company like Netflix. These connections imply a business relationship between the video streaming company and the packet transport service and thus invite additional business issues, such as pricing using two-sided market principles and disputes over discrimination. The combined artifact is an actual segment of video viewed by a user, as illustrated in Figure 4. The dashed line in the figure represents the requirement of ongoing connection between the facilities and artifacts.

In both the spreadsheet and Netflix examples, the technologies can be broken down using the technology flow framework. The main difference is that in the Netflix example the artifacts cannot be combined into useful output without ongoing connection to the facilities. In both examples, there are alternative technologies that change the connection requirements. Online applications like Google Docs make the spreadsheet into a packetized service that is quite similar to a Netflix movie. Downloadable video creates a video file artifact that can function separately from the facility that created it and separately from the Internet.

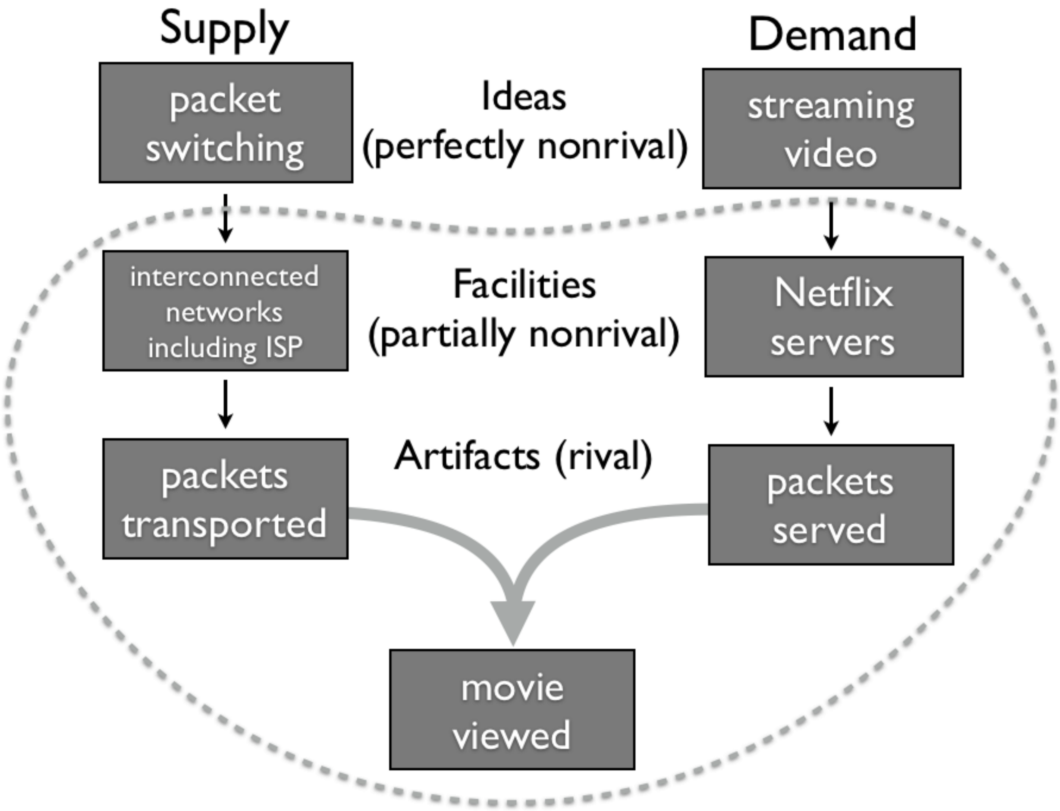

Figure 4: The Internet and one use, Netflix video

5. Infrastructure and GPTs: Similar on Demand Side, Different on Supply Side

So what makes infrastructure different from a GPT? Our answer is that on the demand-side, there is not much difference. Both are inputs into a wide range of downstream uses, as their definitions in Frischmann and LCB, among others, make clear. In both cases, the artifacts supplied are then used in a tremendous variety of applications that create some other kind of useful artifact. For example, both the internal combustion engine and electricity produce artifacts (engines and kilowatt hours) that are used in a huge number of downstream technologies to create other useful artifacts. The fact that it seems more natural to call electricity "infrastructure" than the internal combustion engine does not have much to do with the variety of uses available.

The difference lies with the supply-side facility. For many technologies, the facility is a factory. Some inputs used in a factory, such as designs, organization, and large-scale machinery are partially nonrival. Others, such as labor time, materials, and energy are completely rival. The result is that supplying an additional artifact consumes real resources and has a significant marginal cost. In contrast, other technologies have less modular facilities with extremely large economies of scale (often they are networks built to connect many users). The material, energy, and labor inputs to these large-scale facilities are negligible for producing off-peak artifacts. Only when output becomes so high that the facility becomes congested does supplying an additional artifact create a significant marginal cost. This, we believe, is the difference between infrastructures and GPTs. While other types of GPTs are produced in facilities using rivalrous processes, those technologies that are "infrastructure" have facilities which produce artifacts at negligible marginal cost over a substantial range of output – they are partially nonrival. This may cause market imperfections in provision of the facility. There may be barriers to entry, natural monopoly, and large difference between average and marginal costs – all problems familiar in regulatory economics.

We do need to recognize that there are facilities that produce artifacts under conditions of partial nonrivalry but do not have a wide range of uses in the economy. Software code like Excel is an example, as are print newspaper printing plants and public or club swimming pools. The same might be said of a flour mill in a medieval town or a server farm for cloud computing. Sometimes facilities like these are casually referred to as "infrastructure," but this use of the term is imprecise. Perhaps to be complete we should always differentiate between "special purpose infrastructure" and "general purpose infrastructure," but we believe that the best use of the word really only refers to the general purpose case. When special-purpose technologies are called "infrastructure," we think the label is really intended to mean "important" or "essential to the activity." The word "platform" is often used for technologies that are intermediate between highly specialized and very general, and this is probably more precise.

We believe that nonrivalrous use of the supply-side facility is the key differentiator that makes the Internet *both* an infrastructure and a general purpose technology. In contrast, "the computer" is produced in a rivalrous manner like any other manufactured good. So while the computer is certainly a general purpose technology based on its demand-side uses, it is not properly an infrastructure. This says nothing about the relative impact or importance of the Internet versus the computer, only that the facilities have different production properties.

Our conclusion is that nonrivalry is important for GPTs and infrastructure both. For both, there is nonrivalry in the ideas or principles for *use* which causes positive spillovers for the economy. But in infrastructure there is also a substantial degree of nonrivalry in the supply-side *facilities*, and this creates important differences in how infrastructure is provided and regulated.

6. Openness and Tethering

Demand-side uses for both the computer and the Internet are based on completely nonrival ideas. These ideas are also open and nonexcludable except to the extent that they can be protected by a patent. The result for our examples from Section 4 is that there are many competing sources of spreadsheet programs and of streaming video content, each a separate company with its own facility. There is no "manager" with control over how either technology is used or what facilities are built, so the variety of artifacts available is large. Many of the actual facilities that produce these artifacts, however, are private and do have "managers" that regulate their use, and they tend to function effectively in a private market context.

These examples suggest that the source of spillovers to the general economy is the openness of the *ideas* for using the technology. By having many ideas, and allowing each idea to be reused by many agents, the range of uses of the technology rises and their value to the overall economy multiplies. The facilities and artifacts that actualize these ideas do not need to be open.

The large range of uses is closely related to the concept of indirect network effects. As more users adopt a technology, more services (uses) will be offered by third parties. For example, the more households have broadband, the more different types of streaming video will be offered.

Both network effects and spillovers can lead to market failure. There can be too little use of the GPT or infrastructure, or there can be too little standardization or compatibility between different elements within the GPT or infrastructure, or there can be mis-optimization for a narrower range of uses than is socially optimal. This means that coordination can be desirable, and with it comes the potential for commons management.

We have seen that infrastructure facilities usually have huge economies of scale that make marginal cost negligible when congestion is not taking place. In most cases, the way such large economies of scale exist is that infrastructure operates as a network connecting various users. Sometimes this is explicit, as in the road and telephone networks that physically terminate at driveways and household jacks. Sometimes it is more virtual, as in a stock exchange that provides a physical space for stock trading or a wireless network that provides communication without a physical connection. In either case, networking is a defining feature of infrastructure that is not present in GPTs made of stand-alone components such as the steam engine or computer.

In these cases where artifacts can only be obtained via an ongoing connection to the supply facility, the service is *tethered* (Zittrain 2008). Zittrain defines tethered services as those which are "centrally controlled." This is crucially important because it creates an ongoing relationship, often with legal ramifications, between the supply-side facilities provider and the demand-side user. In a stand-alone GPT, like the internal combustion engine, the maker of the engine would find it difficult to prevent its use in, say, a tractor instead of a truck. Likewise, makers of computers cannot easily prevent them from being used for, say, electronic banking instead of word processing. With infrastructure, the connected nature makes this type of usage control much more feasible. For example, if the telephone provider does not want the user to contact a dial-up ISP, it can simply block such service unless there is regulation to the contrary. Another example is a toll highway; these are usually required to serve all vehicles, but without regulation a toll road could easily discriminate against trucks of a certain company. Indeed, technology providers may purposely try to tether their facilities to their artifacts – Zittrain calls this "appliancizing."

Liebowitz and Margolis (1994, pp. 135–136) noted this difference and used the term "literal networks" to describe what we call tethering: "In some networks, participants are literally connected to each other in some fashion. The telephone system is one such network, as are pipeline, telex, electrical, and cable television systems. These 'literal networks' require an investment of capital, and there is a physical manifestation of the network in the form of pipelines, cables, transmitters, and so on. It is not only feasible but almost inevitable for property rights to be established for these types of networks. Those who attach to such networks without permission from the owner, or who attach without adhering to the rules, may be disconnected, a characteristic that removes the problem of nonexclusion."

We subdivide tethering into three categories:

physical tethering – the facility physically connects with those who use the artifacts (broadband networks, road networks)

virtual tethering – the facility requires (or is engineered to require) a non-physical connection with those who use the artifacts (wireless phone and data, iOS App Store)

legal tethering – a license or other legal permission is required to use the facility and obtain artifacts (patent licenses, software licenses, permits for use of public lands)

Because of the difference in the prevalence of tethering, infrastructure is usually "managed" by its provider, while GPTs in general may or may not be. Theories of GPTs emphasize the interdependence of the GPT itself and the many components that work with it. The need to create and coordinate many components leads to games that are characteristic dilemmas of commons management. There are often prisoner's dilemma games in which component makers free-ride and do not produce enough components and coordination games where component makers need to agree on standards and interfaces.

These problems suggest that GPTs have features that might make commons management desirable. However, GPTs like the steam engine or biotech experience untethered diffusion, and it is difficult to imagine "managing" them or preventing "permissionless innovation."[7,8] The tethered diffusion of infrastructure GPTs like railroads or the Internet, on the other hand, makes them much more "manageable." In particular tethering means that two of Ostrom's (1990) "design principles" for commons management apply more easily to infrastructure than to GPTs: clearly defined boundaries and monitoring.

Infrastructure has more "clearly defined boundaries" than other types of GPTs. Because infrastructure is tethered, it is much more clear who is and is not using the infrastructure. Generally those who wish to use infrastructure must formally become customers of the infrastructure provider, or go through some related process such as registering their car in order to drive on public roads. The owner of a tethered technology retains "bouncer's

7 This term seems to have been first used by Vint Cerf, David Reed, Stephen Crocker, Lauren Weinstein and Daniel Lynch in an Oct. 2009 letter to then FCC Chairman Julius Genochowski supporting network neutrality: http://voices.washingtonpost.com/posttech/Net%20Pioneers%20Letter%20to%20Chairman%20Genachowski%20Oct09.pdf. Thanks to Henrik Rood for noting this connection and source.

8 Increasingly there are field-of-use restrictions in licensing agreements that legally "tether" the technology to certain uses (Schuett 2012).

rights" (Strahilovetz 2006) to prevent use. Boudreau and Hagiu (2009, pg. 7) note this and then add, "The power to exclude also naturally implies the power to set the terms of access (e.g. through licensing agreements)—and thus to play a role somewhat analogous to the public regulator."

Second, again due to tethering, infrastructure usage can be more easily monitored. Most infrastructure systems are purposely set up with monitoring technology to prevent overuse or misuse, to facilitate billing, and increasingly to collect data about users and their activities. Ostrom suggests that for successful commons management, it is important that monitors are accountable to appropriators, and that appropriators design the rules with monitoring costs in mind. It is not clear, however, that many infrastructures are set up this way. By contrast, GPTs are almost impossible to monitor, except through patent licensing, and even monitoring the standards around a GPT is hard.

7. Implications

Policies regarding infrastructure are complicated because of the many linkages between the infrastructure and other sectors of the economy (Hogendorn 2012). Aghion et al. (2009) discuss the difficulty of linking theoretical approaches to real-world policies that promote science, technology, innovation, and growth systems (STIGS). They note the "GPT rationale" for public investment, but also caution that policy responses are more difficult when there are many complex complementarities rather than isolated market failures. Bauer (2014, pg. 671) emphasizes that it is incumbent upon the "analyst to take the relevant interdependencies among players connected by a platform into account." Our technology flow framework puts structure on this: infrastructure policy that affects *artifacts* only affects downstream factor prices, whereas policy that affects the *facility* may actually change downstream production functions or lead to the exit or entry of goods.

Tethering is the reason why the openness question tends to revolve around infrastructure but not other types of general purpose technologies. Openness to innovation is equally important for all of them, but tethering creates a situation where innovative use often requires permission from infrastructure facilities providers. Thus, it is not surprising that computers are generally not subject to much government regulation whereas Internet Service Providers increasingly are.

This comment also gives our answer to Field's (2008) criticism of the broad versus narrow definition of the GPT. Since our framework does not differentiate between how closely integrated the use is with the supply, we are clearly adopting the broad definition. Thus, we have no problem calling both the steam engine and the railway general purpose technologies since both led to a wide variety of uses. The fact that one major use of the steam engine was the railway does not detract from the generalness of either technology. We add that railways are infrastructure as well as GPT because railways have partially nonrival, tethered facilities, but the artifact produced by railways – freight shipped – is indeed used throughout the economy.

A great deal of research has tried to measure the impact of ICT on the economy. It generally follows on Paul David's seminal "The Dynamo and the Computer" (1990) which gives two GPTs in its title and keeps them parallel by choosing the device that serves as an engine rather than any system that goes with it. Jovanovic and Rousseau (2005) are also clear that by "IT" they mean the microprocessor and personal computer. But most recent papers expand from the computer to the whole category of ICT – see the survey by Cardona et al. (2013).

While the contribution of the ICT sector to the economy is a worthy question, "ICT" cannot be cinched into the technology flow framework. There is no single artifact which can be considered the output of ICT, and ICT as a whole is not "a single generic technology, recognizable as such over its whole lifetime" as LCB put it. This suggests that the computer and the Internet are two separate GPTs, just as in LCB's list.

Empirical work on GPTs frequently searches for evidence of spillovers from the adoption of a GPT, either in R&D productivity or in total factor productivity (TFP)growth. Some authors use the strength of spillovers to *test* whether a technology can be labelled a GPT, for example Feldman and Yoon (2011) and Liao et al. (2016). Others use the definition to declare a technology a GPT prior to further analysis, for example Jovanovic and Rousseau (2005). We believe our framework helps with either approach, either to define a GPT or to select a candidate GPT that will be subject to empirical testing. But in both approaches, it suggests that spillovers should be measured across supply-demand sectors that can be modeled using the technology flow framework. If this is taken to mean that ICT should be separated into computers and the Internet (or possibly other sectors such as content), this would reframe the question of spillovers across sectors.

We believe that the most important part of this reframing would be to change the view of spillovers that occur *within* the category ICT as defined by statistical authorities. Consider that computers and the Internet are both categorized in this broad sector of the economy, whereas steam engines and railroads were categorized as separate sectors. Yet the contribution of computers to the Internet seems entirely comparable to the contribution of the steam engine to railroads.

This might alter the research questions around spillovers within ICT. For example, imagine a country that did not produce computers but had a healthy Internet sector based on imported computers. This would depend very much on the degree of tethering and the permissiveness of the innovation environment. The move from the general purpose PC to the more appliancized phone and tablet could affect the hypothetical country's tech sector. Venturini (2015) found that OECD countries received high TFP spillovers from domestic production of ICT goods but not from importing ICT goods. Perhaps this is because imported technology remains tethered to foreign firms and thus cannot serve as an innovation platform to create downstream spillovers in the importing country.

8. Conclusion

We have discussed two important terms, general purpose technology (GPT) and infrastructure. The two are sometimes the same and sometimes different, and the exact relationship has not been clear in the literature. Here we argue that the key is to split up both kinds of technology into their demand-side and supply-side properties. On the demand side, the two are essentially the same, which explains why they are often lumped together when discussing their impact on the general economy. But on the supply side, infrastructure has partially nonrival production facilities that make it more likely to be subject to regulation and also more likely to employ commons management techniques. In particular, the fact that infrastructure is usually built as a network and involves some *tethering* of the user to the provider means that maintaining boundaries and monitoring use is much easier in the context of infrastructure. This in turn explains why the openness principle becomes more important in the infrastructure context.

We hope our technology flow framework will be useful in two ways. First, it provides guidance on how and where to look for positive spillovers that are often relevant in regulatory and antitrust proceedings. Second, it clarifies the use of the terms infrastructure and general purpose technology which are themselves often used to bolster contentious claims in policy debates. We also hope that it will help spur further research on infrastructure as a microfoundation of economic growth.

References

Aghion, P., David, P. A., & Foray, D. (2009). Science, technology and innovation for economic growth: linking policy research and practice in 'STIG Systems'. *Research Policy*, *38*(4), 681-693.

Bauer, J. M. (2014). Platforms, systems competition, and innovation: reassessing the foundations of communications policy. *Telecommunications Policy,* 38(8), 662–673.

Boudreau, K. & Hagiu, A., 2009. Platforms Rules: Multi-sided Platforms as Regulators. In A. Gawer, ed. *Platforms, Markets and Innovation.* London: Edward Elgar.

Bresnahan, T. (2010). "General purpose technologies," Chapter 2 in *Handbook of the Economics of Innovation*, 761–791.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'?. *Journal of Econometrics*, *65*(1), 83–108.

Cardona, M., Kretschmer, T., & Strobel, T. (2013). ICT and productivity: conclusions from the empirical literature. *Information Economics and Policy* 25(3), 109–125.

Paul A. David, (1990) "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Review,* Vol. 80, No. 2, pp. 355–361.

FCC, *2010 Open Internet Order,* FCC 10-201, December 23, 2010.

FCC, *2015 Open Internet Order,* FCC 15-24, March 12, 2015.

Feldman, M. P., & Yoon, J. W. (2012). An empirical test for general purpose technology: an examination of the Cohen–Boyer rDNA technology. *Industrial and Corporate Change*, *21*(2), 249-275.

Field, A. (2008). Does Economic History Need GPTs?. *Available at SSRN 1275023*.

Frischmann, Brett M. *Infrastructure: The Social Value of Shared Resources*. Oxford University Press, 2012.

Frischmann, B. M., & Hogendorn, C. (2015). Retrospectives: The Marginal Cost Controversy. *Journal of Economic Perspectives*, 29(1), 193-205.

Helpman, E., & Trajtenberg, M. (1998). Diffusion of General Purpose Technologies. *General Purpose Technologies and Economic Growth*, 85.

Hess, C., and E. Ostrom. Ideas, Artifacts, and Facilities: Information as a Common-pool Resource. *Law and Contemporary Problems* 66, no. 1/2 (2003): 111–145.

Hogendorn, C. (2005). Regulating vertical integration in broadband: open access versus common carriage. *Review of Network Economics*, *4*(1).

Hogendorn, C. (2012). Spillovers and network neutrality. Chapter 8 in *Regulation and the Performance of Communication and Information Networks*, 191.

Jovanovic, B., & Rousseau, P. L. (2005). "General Purpose Technologies" Chapter 18 in *Handbook of Economic Growth,* Volume 1, 1181–1224.

Liao, H., Wang, B., Li, B., & Weyman-Jones, T. (2016). ICT as a general-purpose technology: The productivity of ICT in the United States revisited. *Information Economics and Policy.*

Liebowitz, S. J., & Margolis, S. E. (1994). Network externality: An uncommon tragedy. The Journal of Economic Perspectives, 8(2), 133-150.

Lipsey, R. G., Carlaw, K. I., & Bekar, C. T. (2005). *Economic transformations: General purpose technologies and long term economic growth.* Oxford University Press.

Majumdar, S. K., Carare, O., & Chang, H. (2010). Broadband adoption and firm productivity: evaluating the benefits of general purpose technology. *Industrial and Corporate Change*, *19*(3), 641-674.

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1989). Industrialization and the Big Push. *The Journal of Political Economy*, *97*(5), 1003-1026.

Ostrom, E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge University Press.

Ramsdell , Robert E (November 1980). "The Power of VisiCalc". BYTE. pp. 190–192.

Rosenstein-Rodan, P. N. (1943). Problems of industrialisation of eastern and south-eastern Europe. *The Economic Journal*, *53*(210/211), 202-211.

Schuett, F. (2012). Field-of-use restrictions in licensing agreements. *International Journal of Industrial Organization,* 30(5), 403–416.

Strahilovetz, L. (2006), 'Information asymmetries and the rights to exclude', *Michigan Law Review,* 104 (8), August, 1835–98.

Venturini, F. (2015). The modern drivers of productivity. *Research Policy,* 44(2), 357–369.

Zittrain, J. (2008). *The Future of the Internet–and How to Stop It.* Yale University Press.