Aligning modes of organization with technology: Critical transactions in the reform of infrastructures

Rolf Kunneke, John Groenewegen, Claude Ménard

Aligning modes of organization with technology: Critical transactions in the reform of infrastructures

Rolf Künneke, John Groenewegen, and Claude Ménard*

JOURNAL OF ECONOMIC BEHAVIOR AND ORGANIZATION

75 (3): 494-505, 2010

Corresponding author: Rolf Künneke, Economics of Infrastructures, Faculty of Technology Policy and Management, Delft University of Technology, P.O. Box 5015, 2600 GA Delft, The Netherlands, Tel: +31 (0) 15 2787752, E: R.W.Kunneke@tudelft.nl

* Künneke and Groenewegen, Delft University of Technology; Ménard, University of Paris (Panthéon-Sorbonne). Different versions of this paper were presented at various seminars and workshops in Europe and abroad. We are grateful to participants and discussants for their comments and critiques. We are particularly indebted to Richard Green, Carine Staropoli, and two anonymous referees for helpful comments and suggestions.

Abstract

This paper is about the alignment of technology and modes of organization in infrastructures in the context of their reform. Since infrastructures are characterized by strong technical complementarities, we explore the resulting 'critical technical functions' that need to be performed in order to guarantee the expected technical performance of the system. We characterize 'critical transactions' as essential to provide adequate support to these functions. We distinguish various modes of organization that can effectively coordinate these critical transactions. We argue that the features of these transactions determine the alignment between organization and technology and should be taken explicitly into account when reforming infrastructures.

Key words: Transaction costs, technology, institutional change, reforms of infrastructures

I. Introduction

Reforming infrastructures has been high on the agenda of economists and policy makers over the last three decades. Introducing 'more market' has been at the core of these reforms, with the underlying rationale that this is the adequate way to increase efficiency, to stimulate suppliers to serve consumers' needs, and to innovate in products and processes. In addition changes in the allocation of property rights, with the reintroduction of private interests in infrastructures, have often been considered inevitable in order to finance the huge investment needs in sectors like telecom, energy, rail transport, and water. For example, it has been estimated that investments of a magnitude of over $ 100 billion (in 2003 US dollars) within a decade would be required to halve the percentage of population without access to drinking water and basic sanitation.[1]

Although changes in property rights and the related decision rights may take many different directions, almost all restructuring of very diverse infrastructures seems to follow a similar recipe. Infrastructure firms that have historically developed as vertically integrated monopolies are separated into different entities in order to allow for competition. Monopolistic activities (exploitation of network related services) are unbundled from activities for which it is assumed that competition can be introduced (production and supply of different services). The monopolistic segment needs to be closely regulated in order to safeguard a level playing field. For this purpose independent regulatory agencies are established that typically supervise the access and pricing of the network services. Simultaneously, particularly with respect to the competitive segment, there is a switch from a command-and-control approach to a contractual one.

[1] UN Millennium Project (2005).

Following this pattern, liberalization seems essentially a matter of institutional change, i.e. the design of well functioning markets in order to provide sufficient incentives to serve the above mentioned objectives. Technology is not considered an important factor enabling or frustrating this process of change. Most reforms of infrastructures assume that technology will change autonomously and adapt quite smoothly, thanks to the incentives provided by the new rules of the game.

In this paper, we argue that infrastructures are complex technical systems with a strong degree of complementarity that is enabled through network relationships. Accordingly, all major elements technically interact in a specific manner with each other, which makes it possible to produce certain services like the provision of energy, telecommunication, clean drinking water, or rail transport. These complex technical systems have very peculiar economic features including path dependence, lock-in effects, multiple and possibly inefficient equilibria.[2] Hence, it might be expected that the technological status of various infrastructures influences the opportunities for restructuring and contributes to shape the resulting performance. Building upon the literature on co-evolution,[3] with a view at how transactions are organized or can be organized in that context, we aim to further elaborate on the mutual relations between institutions and technology in infrastructures. Focusing mostly on one aspect of performance, the technical one, we show that the technical functioning of infrastructures needs to be supported by appropriate modes of organization. Insufficient or failing institutional support results in inferior technical performance contributing to unreliable services or in extreme cases even

[2] David (1985), Arthur (1988), Economides (1996).

[3] See, for instance, Nelson (1994), Dosi (1982), Von Tunzelmann (2003), Perez (2002), Saviotti (1996), Murmann (2003).

to system outages. There is empirical evidence that the technical functioning of at least some infrastructures did not meet expectations after reforms, particularly when reforms combined unbundling and changes in the allocation of both property rights and decision rights. Refer for instance to the various blackouts in electricity systems[4], or some train accidents that followed reform of the British rail sector.[5] We take such anecdotal evidence only as an incentive to better understand the role of technology in the process of regulatory reforms in infrastructures. What we intend to do is to identify critical technical functions and to show how they impose properties on transactions that should be reflected in the adequate alignment of the technical functions with the modes of organization. Hence this article takes a theoretical perspective and does not intend to further explore the incidence of specific technical malfunction in various infrastructures.

The core of our argument is that the technical complementarity between the various elements of infrastructure systems most of the time imposes a technical need for coordination with respect to functions like capacity management, system management, interconnection, and interoperability.[6] These technical functions can be considered as **critical** for safeguarding the technical performance of infrastructures. In the railway sector,

[4] For historical trends of large blackouts in North America between 1984 and 2006 see Hines, Apt and Talukdar (2009). We argue in this paper that in times of technical distress the institutional support of critical technical functions is essential to safeguard the functioning of the infrastructure system.

[5] Gourvish (2008) provides a comprehensive historical analysis of the reform of British railways between 1997 and 2005. The Hatfield crash on October 17th, 2000 (four people were killed and 70 injured), was a watershed in the reform of this sector from privatization and liberalization back to more stringent governmental monitoring and control. We interpret this as a possible indication that certain critical technical functions are not supported by sufficiently effective modes of organization.

[6] Finger, Groenewegen and Künneke (2005), p.240 – 241.

for instance, traffic control is very crucial to maintain the technical reliability of the system. A failure of this aspect of capacity management can cause accidents and threaten the technical integrity of the infrastructure. Similar examples are the monitoring of the quality of drinking water (poor quality might cause severe health problems) and load balancing in the electricity sector (blackouts occur if the system is imbalanced).[7] Taking this technical criticality as point of departure, the question arises whether we can identify supporting 'critical transactions' that are essential for the functioning of infrastructures. In other words: if critical transactions are not well coordinated, some technological critical functions of the systems are not satisfied so that the system severely fails to deliver the expected services. We are interested in modes of organization that guarantee the coordination of critical transactions related to critical technical functions of liberalized infrastructures.

In order to explore this issue, we are building on different strands of literature. New Institutional Economics provides an important source of inspiration with its emphasis on the alignment of modes of organization with transactions at stake (Williamson, 1996) as well as with the identification of the key role played by institutional factors in technological changes (North, 1990 chap.9) and in the reform of infrastructures (Joskow, 1997). However, as acknowledged repeatedly by Williamson, the transaction cost perspective that provides foundations to this approach has not really integrated the technological dimension.

Our paper intends to make a step in this direction. We perceive infrastructures as complex socio-technical systems in which technological, economic, political, and social features strongly interact with each other (Perez, 2002, Van de Poel, 2003, Kroes et al.,

[7] Load balancing refers to the technical balancing of production and consumption in electricity networks that needs to be guaranteed at every moment of time.

2006, Geels and Schot, 2007). We focus on problems associated to the technical control of these systems (Nightingale et al., 2003), which we perceive as critical technical functions. Traditionally this is the field of control engineering that deploys very sophisticated technical approaches to secure the proper technical functioning of complex technical systems (Dutton et al., 1997). We are interested in relating technical characteristics of these control mechanisms to different modes of organization. Hence we focus on issues associated with the effectiveness of transactions that govern these technical control mechanisms.

The article is structured as follows. Section II specifies the notion of technical criticality in infrastructures. Critical technical functions are determined from a control engineering perspective, which is an important way to aligning the actual to the desired technical performance. Section III defines critical transactions, which are transactions providing essential support to the critical control mechanisms. Section IV identifies modes of organization to secure these critical transactions. Differences in critical transactions and modes of governance across infrastructures and over time are elaborated in Section V. The main findings are summarized in Section VI.

II. TECHNICAL CRITICALITY IN INFRASTRUCTURES

A core argument that we develop in this paper is that technology imposes critical functions and that the benign neglect of this issue in reforms of infrastructures is reflected in misalignments of modes of organization with the requirements of the critical technical functions. We put special emphasis on control-related problems.

II.1 Criticality: a multifaceted concept

The notion of criticality is fuzzy and hence difficult to define and delineate. Often the term 'critical' is used to qualify things or circumstances considered significant or fundamental, whether the connotation is positive or negative. Refer for instance to expressions like 'critical mass', 'critical information', 'critical rules and regulation', 'critical services' and 'critical infrastructures'. This article does not elaborate on the problem which infrastructures are to be considered critical and for what reason.[8] Rather, we take it as a given that there are infrastructures that are fundamental for the functioning of our societies (e.g., the provision of drinkable water). Therefore, in this paragraph we focus on identifying those aspects of the technical operation and management of any given infrastructure that are critical in order to meet expectations with respect to the technical functioning of these systems. In order to determine criticality in this context, we need to reflect on two aspects:

- How to specify expectations with respect to the technical functioning of infrastructures?
- How to identify those aspects of the technical operation and management of infrastructures to be considered as critical with these expectations in mind?

Expectations with respect to the technical functioning of infrastructures include reliability, safety, and security of supply. Technical **reliability** refers to the ability of the infrastructure system to produce certain services at the expected place within the expected period of time. For instance, electricity has to be available to all users connected to the network at all moments of time. **Safety** deals with the secure use of the services delivered

[8] A report to the US Congress (Moteff, Copeland and Fischer 2003) made such an exhaustive list of components that are 'critical' for the running of an economy, from agriculture to airline to the banking system.

through infrastructures, like for instance clean drinking water or secure means of rail transport. Since infrastructures might cause significant external effects, non users need also to be protected against the possible consequences of the technical malfunctioning of these systems. The safety of nuclear power plants illustrates this point. Finally **security** of supply represents the ability of infrastructures to sustain their activities in the foreseeable future. In the energy sector for instance, primary energy sources need to be available for the production of electricity.

Referring to the second question, the criticality of the operation and management of infrastructures is very much determined by the strong degree of technical system complementarities. Since the components of infrastructures are in one way or another connected through a physical network, they cannot be operated independently from each other. Hence, criticality can occur in at least two different manners. **First**, there might be assets that are fundamental for the functioning of the system (Moteff et al., 2003). Examples of critical assets are traffic control of railroad systems, transmission lines in electricity, or pumping stations in the water sector. Failures of critical assets lead to serious interruptions not only in the infrastructures in question, but throughout society.[9] Obviously, an empirical identification of these critical assets requires a very detailed analysis of different infrastructures, e.g., their specific geographical settings.[10] **Second**, technical

[9] Moteff, Copeland and Fischer (2003) provide an overview for the case of the War Against Terror in the USA.

[10] For instance in the USA, "the electric power infrastructure includes 92,000 electric generating units (including fossil fueled, nuclear, and hydroelectric units), 300,000 miles of transmission lines, and 150 control centers, regulating the flow of electricity. The nation's water infrastructure includes 75,000 dams and reservoirs, thousands of miles of pipes and aqueducts, 168,000 public drinking water facilities, and 16,000 publicly owned waste water treatment facilities". Moteff, Copeland and Fischer (2003), p.8.

criticality can depend on specific functions that are essential to safeguard the technical performance of infrastructures, such as timetabling in the railroad sector or load balancing in electricity. Although these two interpretations of criticality have unambiguously complementarities, here we focus mostly on this second aspect, that is, the technical functions of criticality. In order to do that, we first proceed with a short overview of control issues in an engineering perspective.

II.2: A control engineering perspective

As already mentioned, one of the basic features of infrastructures is the high degree of technical complementarity between its different components. This requires coordination in order to safeguard the satisfactory functioning of the system. In order to facilitate this complementarity, control mechanisms need to be established that monitor important performance parameters. For instance, in railways slots need to be allocated in such a way as to avoid delays or, even worse, accidents.

Control engineering is concerned with the planning and control of complex technical systems. Mathematical modelling is used to analyse the dynamic behaviour of infrastructures and to contribute to their reliable operation and management.[11] An important objective of control engineering is to align the actual performance of a technical system with the desired performance. In electricity, the voltage level needs to be stable within a certain bandwidth. In the water sector, the physical network must meet requirements imposed by the laws of turbulence. On a very general level there are two possibilities to model these kinds of control problems, i.e. the so called open and closed control systems.

[11] See for instance Dutton, Thompson and Barraclough (1997)

Open control systems are based on the assumption that the relation between input[12] and system performance is sufficiently stable to satisfy expectations, while ensuring the technical functioning of the system. The operation of the system is predetermined and there is no feedback between the actual and the desired performance.[13] In infrastructures, the use of public roads is an example. The availability of certain types of roads (i.e. the systems input) is assumed to sufficiently predict the desired performance (i.e. the accessibility of certain places in a certain time). Traditionally there is no feedback between the desired and the actual time of travel.[14] The technical management and operation of these open control systems is not directly concerned with actual use and performance, but primarily deals with the proper provision of the required production factors (i.e. the input). This case is rather the exception than the rule in infrastructures. We therefore concentrate on the more complex case of closed control systems.

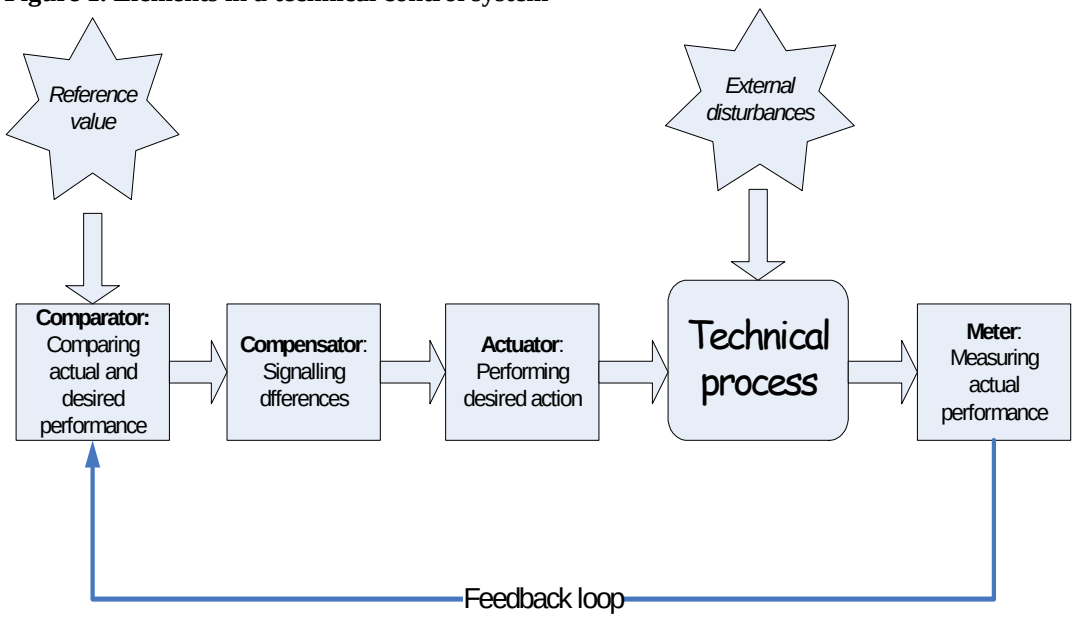
Closed control systems are characterized by a feedback loop between the actual and desired performance and the ability to adjust system performance in case of intolerable differences between both. Examples of these closed loop control systems can be found in many infrastructures, as with the allocation of scarce network capacity in electricity, railroad or water infrastructures. Elements of such technical control systems are summarized in (Figure 1).

[12] In control engineering the term input refers to the production factors (for instance a power plant, primary energy sources, and labor) that are involved in a complex system.

[13] The classical example of an open loop system is a washing machine. The predetermined programs are operated without actual feedback of the users whether they are satisfied with the result or not.

[14] Recently more advanced dynamic systems of road traffic control have been developed in order to cope with the increasing problems of congestion. These can be considered as closed loop systems.

Figure 1. Elements in a technical control system



The core element of this scheme is a technical process that is subject to external disturbances influencing its performance. In the electricity sector, for instance, the demand for power fluctuates even within very short time periods, because users switch their electrical appliances on and off in a not completely predictable way. Since there are no opportunities for storage, the power production needs to be adapted instantaneously in order to meet this changing demand and thus safeguard the technical reliability of the power system. Figure 1 depicts various technical interventions necessary to accommodate a feedback between the actual and desired performance. Starting from the left to the right the technical elements of a closed system can be characterized as follows. The *comparator* compares the actual performance with the desired performance. The desired performance is assumed to be externally given as a *set point or reference variable*. The system performance is operationalized in terms of proxy variables that are measurable with a sufficient degree of accuracy. Important technical performance parameters of the electricity system include the voltage level (for instance 240 Volt) and the frequency (60 Hertz). Possible deviations are signalled to the *compensator*, which is a device that calculates the

appropriate action to be taken. A dropping frequency of alternate electric current can be compensated by increasing the spinning reserve of electricity generation. The *actuator* performs this desired action. Finally the technical system is adjusted as a reaction to the external disturbances. In the next control cycle, the actual performance is measured again, and a new feedback loop is initiated.

Control engineering contributes to the stability and robustness of infrastructures.[15] After an external disturbance the system needs to return to a technical equilibrium position. The time to respond is an important feature for at least two reasons. First, the costs associated to the unsatisfactory functioning of the system need to be minimized. These include the costs of the actuation (i.e. correction) of the disturbance and the costs imposed to the system as a consequence of the malfunctioning. Obviously there is a trade-off between both. Second, certain technical processes imply a time span of control that cannot be exceeded without threatening the system integrity.[16] With respect to operation and management, these processes typically require a short-term reaction. An extreme case is the load management between electricity production and consumption, which demands an almost instantaneous response of only micro- seconds.

Technical control engineering is typically concerned with the short term balancing of the day-to-day operation of complex systems, usually under very rigid assumptions. The production process is assumed to be stationary with steady-state dynamics. Often there is

[15] Robustness can be defined as 'the degree to which a system or component can function correctly in the presence of invalid inputs or stressful environment conditions'. Institute of Electrical and Electronics Engineers (1990).

[16] Technical system integrity describes a situation in which the technical performance of the system meets expectations.

only one parameter for the desired output, and one parameter for controlling the system, i.e. 'single input single output' (SISO) systems. Even under these conditions, control system modelling is a very challenging task.[17]

Of course other, more long-term oriented factors matter that partially condition technical reliability and long term security of supply, e.g., adequate investments or appropriate maintenance of the system. In our analysis we deal with control issues varying from very short term to long term. In this context, the representation provided by the analysis of closed control systems is particularly helpful in identifying critical technical functions of infrastructures.

II.3: Critical control in infrastructures

If we match our definition of criticality with the characteristics that closed control systems point to, control mechanisms can be labelled as critical if:

(1) They imply a sufficient *technical scope of control* and they are *unique* in the sense that there are no alternative control mechanisms that perform similar tasks. Hence, they are crucial for the reliable technical functioning of the infrastructure system. A failure will have system wide consequences.

(2) They involve strong *time constraints*, since critical control mechanisms need to be activated in a specific and often very short period of time.

An important dimension that interferes with critical technical control mechanisms in infrastructures has to do with capacity management of network related resources. Indeed,

[17] 'Multiple input multiple output' (MIMO) systems of control are sometimes estimated by a linear combination of several SISO models. Dutton, Thompson and Barraclough (1997), p.20

network capacities are usually limited and need to be allocated for different users in different periods of time for various objectives. In the railway sector the time slots need to be allocated in a way that allows organizing a timetable which meets travellers' expectations for timely and fast relocation between the various nodes in the network. In the drinking water sector the pressure in the pipeline systems needs to be stabilized within the constraints imposed by the physical capacities of the network in order to meet the needs of the consumers with continuity.

There is a discussion in control engineering about whether it is possible to avoid short term control problems by a more intelligent design of the technical system. Short term control problems might be mitigated by a thorough long term planning. However, there are limits to this approach. Complex technical systems typically evolve over time and control problems arise as the system is evolving, without an integrated technical planning.[18] For instance, the electricity system evolved from a local system, to regional, national and supranational – an evolutionary development without any overall blueprint. In water systems, the changing demography of cities as well as changes in architectural design makes long term technical planning also difficult and requires adaptability in control mechanisms at a level which is hard to anticipate.

III. CRITICAL TRANSACTIONS

The technical criticality of some functions in the adequate provision of infrastructures and the conditions they impose on control mechanisms have an impact on how transactions should be organized among different nodes of the system and on the resulting costs. In this section, we focus on this transactional issue, i.e., on transactions that

[18] Dutton, Thompson and Barraclough (1997), p.2

are critical in the provision of supports for control mechanisms in infrastructures. Transactions are about the transfer of 'rights to use' goods or services across technologically separable interfaces (Williamson, 1985, Prologue, p. 1; Ménard, 2005 p. 284). Two aspects are particularly significant in the concept of transaction thus defined. (1) The very existence of transactions requires that activities can be separable from a technological point of view. If technically this is not possible, there is no choice in terms of organization. However, if separability is feasible, the issue of how to organize transactions emerges: alternative solutions must be considered. (2) Separability from an economic viewpoint introduces the possibility of defining 'rights' on assets needed and of transferring these rights. There are different ways to do so, which introduces the problem of alternative (and competing) modes of organization.[19] For example, different railroad companies can use the same rail track system, so that rights of access, which are typically rights to use, must be defined and allocated. Selling train tickets can be integrated as one activity among others in a railroad company, or can be outsourced to an independent company, which raises the issue of the organizational arrangement.

Hence, not all transactions have the same status, in that not all transactions are related to critical control mechanisms. For instance, trains with rights of access must be monitored by a traffic control centre in order to guarantee safety, while they can very well run without an entity selling tickets (e.g., public authorities can decide that using public transportation is free). We call 'critical transactions' those transactions essential to accommodate critical control mechanisms. A failure in 'critical transactions' threaten the capacity of the system to maintain one of the critical control mechanisms defined in the

[19] There is another consequence that we do not discuss here: since rights need to be defined and enforced, adequate institutional supports are essential for the system to function.

previous section. This means that critical transactions require an adequate mode of governance to coordinate critical technical functions. In that respect, they determine the alignment (or misalignment) of the technical dimension of control mechanisms, which in turn determines their criticality, with their governance dimension, which corresponds to allocations of rights that shape alternative modes of organization. This interrelation between critical technical functions and critical transactions is not addressed in the standard contract theory.[20]

In order to capture analytically this interaction, we distinguish two dimensions that shape critical transactions, combining with varying importance: a technical-control dimension, which relates to the scope of control and the speed of adjustment of technical functions;[21] and an organization-specific dimension, which concerns the organizational needs to meet the technical system requirements.

III.1: Technical-control dimension

Critical transactions are essential for guaranteeing the technical functioning of the system, thus imposing constraints on the mode of organization. Let us take the example of interconnection of different networks. Water could not be delivered to households without adequate control of the pressure in the interconnected pipe system. Trains could not interconnect if the rail system differs in width or in the type of energy required. Electricity networks could not function as an interconnected system without load balancing between production and demand of electric power. In other terms, critical technical functions impose

[20] See Salanié (2005)

[21] See our sub-section IV.1 for a definition and discussion of these characteristics.

certain characteristics to the organization of relevant transactions, usually pushing towards integration or towards tight control through a 'strategic centre'. If there are competing producers of electricity, their access to the grid must be coordinated through a complex set of transactions for which the different possible arrangements are limited and all require unified decision rights at some nodes. These 'critical transactions' are associated to the monitoring and control of interfaces in the system, such as turning on and off key facilities, monitoring bottlenecks, and so forth.

The coordination requirements depend on the characteristics of the critical technical functions, i.e. the scope of control and the speed of adjustment. On a very general level it can be stated that the coordination requirements become more stringent the larger the technical scope of control and the shorter the time period to react on disturbances. The case of load balancing in electricity systems illustrates: it has to be performed within micro-seconds on a system wide technical scope of control. Usually this requires a system operator with short term directive decision rights in order to restore the balance of the system. Another extreme is the case of coordination requirements with respect to components of infrastructures for which a long period of time is needed for its realization. Refer for instance to the development of tracks for innovative high speed train or the construction of new international gas pipelines. These rail tracks or pipelines must fit into the existing network and they usually require years for their technical specificities to be designed and implemented. Under these conditions, the need for coordination can be accomplished through different modes of governance, including the implementation of standardization, different types of contracting, or even market competition for the best suitable technical solution.

III.2: Organization-specific dimension

Beside the effects on transactions due to technical factors of criticality, there are indeed critical effects related to the way the economic viability of the system is organized. For example, if a contract does not determine prices guaranteeing adequate return on capital, underinvestment follows, eventually to the point the system collapses. In other terms, there are organization-specific effects which play as essential a role as those identified in the technical-control dimension and must be dealt with in order to assure performance that allow the system to work.

First, highly specific investments that infrastructures require must be made on critical assets of the system, e.g. power plants in electricity or pumping stations in water. We have learned from transaction costs economics that these investments will be made only if transactions are organized in a way that provides safeguards against risks of opportunism, usually under the form of integration or long term contracts.[22] Second, some transactions may face circumstances generating significant uncertainty. The probability of snowstorms affecting the railroad system or transmission lines of electricity in Canada imposes not only technical constraints, but also organizational ones, e.g., capacities to restore the technical functioning of the system within reasonable delays. Choices by decision-makers, whether they are public or private, may also generate different levels of uncertainty: a nuclear plant imposes security measures, and therefore organizational arrangements, radically different from plants using gas propelled turbines. Third, the identification of critical transactions must take into consideration strategic behaviors: some parties may have an interest in defending that a transaction is technically critical because this preserves their monopolistic position. Last, but not least, because of their criticality, the

[22] See Shelanski and Klein (1995) and Klein (2005) for surveys on this issue.

key technical functions of control mechanisms require room for particularly powerful incentives, so as to attract and motivate adequately the highly specific human capabilities required. In that respect, contracts with engineers in charge of the safety in a nuclear plant will have different characteristics than those with engineers working in a coal plant.

To sum up, both technical-control and organization-specific dimensions of critical transactions should align with critical technical functions, pushing either towards integration, so that a unified organization secures the requirements of criticality, or towards hybrid arrangements, with a strategic center coordinating parties that are simultaneously complementing each other and competing. As a result, complex modes of organization emerge.

IV. MODES OF ORGANIZATION TO SUPPORT CRITICAL TRANSACTIONS

In this section we develop different categories of critical transactions by further operationalizing its technical-control dimensions and organization-specific dimensions. This allows us identifying modes of organization that support critical transactions, and thus help to align some technological requirements of infrastructures with appropriate modes of organization.

IV.1: Categories of critical transactions

As mentioned earlier, the technical-control dimension relates to the scope of control and the speed of adjustment. ‘Scope of control’ can be roughly decomposed in at least three different levels. The most extensive scope entails the ‘system’ level, which can be delineated by the most far reaching technical complementarities between its elements, i.e. the nodes and links of the network. In the case of railroad systems, this would entail the

interconnected and complementary activities including traffic control, slots allocation, timetabling, norms and standards with respect to compatibility of equipment, and so forth. A second level could be described as 'subsystem', which constitutes a technically separable part of the overall system. Subsystems are able to perform certain services independently from the overall system, but they are interconnected with it through a technical gateway (or backbone) in order to contribute to system wide services. For instance the railroad system can be divided into long distance and regional services. Regional systems are able to provide regional transport services independently and autonomously. However, for long distance travelling, interconnections to long distance railway tracks and corresponding regional networks are necessary. In this example we can identify the interconnected railway system as the relevant 'system' level of analysis, whereas the regional networks constitute the 'subsystems'. Finally, at the lowest level of abstraction the various 'components' of infrastructures can be identified. In the case of railway systems this would include components like locomotives with adequate propulsion, marshalling yards, and railway stations.

'Speed of adjustment' is interpreted as the time span that is available to perform the feedback loop in closed control systems as indicated in figure 1[23]. This might be measured in time periods ranging from fractions of a second to decennia. Since the exact measurement of time periods is quite specific for each infrastructure sector, we propose to define this time dimension in terms of different degrees of freedom with respect to the completion of the feedback loop (Figure 1). Although time is continuous, for analytical purposes we take four discrete points that have sufficiently distinct characteristics in terms of the operation and planning of the control system:

[23] Joskow (2003) stresses the importance of this point with respect to the electricity sector.

- **T_0:** Operational balancing of the control system, under the condition that all technical and organizational parameters of the system are given, except for the external disturbances. These are typically very short term operations, like load balancing of the electricity system, or traffic control in railways or aviation.
- **T_5:** Capacity utilization under the same assumptions as in T_0, but in addition the operation of the technical process can be altered or modified. For instance, electricity producers can make a choice whether to produce power or not, depending on the price in the spot market. These decisions are typically made within a few hours or days. The existing capacity can be adapted to the specific technical and economic conditions of the system and its users.
- **T_{15}:** Capacity allocation under the same assumptions as in T_5, but in addition the resources that are needed to facilitate the technical process can be allocated. Examples are the allocation of slots for passenger- and cargo transport in railway systems, the allocation of cross border transmission capacity in energy, or slot allocation on airports in aviation. In other words, the capacity can be modified in response to changing needs of users and producers.
- **T_{50}**: System innovation and/ or transformation under the same assumptions as in T_{15}, but in addition the reference values with respect to the expected technical performance of the system are changing. This requires the design of a new control loop, in which all elements, including the technical process, are redesigned. Typically this is a long term process that takes several decades to be effectuated. Examples include system expansion or system modernization, like

for instance the construction of new international gas pipelines, new high speed rail connections, or the development of intelligent networks in electricity.[24]

The operationalization of the technical-control into three values depicting the scope of control and four values describing the speed of adjustment allows combining them into a matrix with twelve different combinations representing different categories of critical transactions (Figure 2). As we demonstrate in the following section, each category represents different organization-specific dimensions as well. Hence, each of these categories imposes different constraints on modes of organization to secure the reliable technical functioning of the infrastructure system.[25]

Figure 2. Modes of organization to secure critical transactions
(Organizational requirements in parenthesis)

[24] Initiative to develop intelligent electricity networks include www.smartgrids.eu (2007) and www.gridwise.org (2007)

[25] In our approach we take the technical-control dimension as the starting point for the identification of critical transactions. This is in line with the objective of our analysis of identifying those transactions that are critical to support the technical functioning of infrastructures. However, other categories of critical transactions might be identified if we would take another starting point of the analysis, like for instance the organization-specific dimension.

Scope of control / Speed of adjustment	**System** *(requires directive intervention)*	**Subsystem** *(requires coordination)*	**Component** *(requires corroboration)*
T_0 Operational balancing *(requires supervision)*	Authoritative supervision ['system operator']	Collaborative supervision ['system regulator']	General framework conditions ['system norms and standards']
T_5 Capacity utilization *(requires monitoring)*	Compulsory monitoring and enforced adjustment	Mutual monitoring and stimulated adjustment	Self monitoring and voluntary adjustment
T_{15} Capacity allocation *(requires facilitation)*	Controlled allocation mechanism	Guided allocation mechanism	Competitive allocation mechanism
T_{50} System transformation and innovation *(requires planning)*	Directive planning	Indicative planning	Decentralized planning

IV.2: Modes of organization to secure critical transactions.

We first discuss modes of organization related to the scope of control (i.e. the columns of Figure 2), followed by the speed of adjustment (i.e. rows of Figure 2). At the system level, critical transactions require *directive interventions* in order to guarantee the technical complementarities that are essential to safeguard the reliability of the infrastructure system. Load balancing in electric power systems or traffic control in aviation and rail transport illustrate what is at stake here. At this level, investments are typically very specific since the associated technical control mechanisms are unique and

represent dedicated assets in the system. Examples include control centres for power systems or traffic flows, which are typically dedicated for their use in specific infrastructure systems. Uncertainty is another important aspect, deriving from the facts that these investments have a very long lifetime and that the technical and economic conditions of the system are difficult to predict. In addition, the extensive scope of control required and its accompanying interdependence create opportunities for strategic behaviour for those who operate and control these facilities. In order to safeguard these critical transactions, strong incentives are needed to attract the necessary resources and knowledge. These organization-specific dimensions pinpoint to modes of organization that centralize decision rights, especially in order to meet the tight time restrictions and to secure the expected system performance.

At the level of subsystems, the degree of technical complementarity is limited to this more or less isolated level; hence assets tend to be less unique. Regional railway networks or local electricity distribution grids provide examples. However, in order to allow the subsystems to technically connect to each other, a certain degree of *coordination* is required. Asset specificity is still an important issue, but since several subsystems with comparable characteristics might coexist, there tend to be opportunities to redeploy investments. Basically, the possible consequences of a technical failure of critical functions are limited to the reliability of the subsystem. Accordingly, from a system perspective the degree of uncertainty potentially decreases as well as the opportunities for opportunistic behaviour. This provides some relieve for the incentives that are needed to safeguard critical functions.

For the components of infrastructures, compliance with the system needs has to be safeguarded. Components need to fulfil technical requirements in order to contribute to the technical complementarity of the system. For example, opening competition to providers of rail freight services require that they use homologated equipment, fitting in the existing capacities. The degree of asset specificity is comparably low, as well as the degree of uncertainty and the possibility for opportunistic behaviour. [26] Modes of organization need to facilitate *corroboration*[27] in order to establish and sustain system complementarities.

Turning to the speed of adjustment, the following requirements with respects to the modes of organization arise. T_0 stands for the operational balancing of infrastructures. In a very short period of time, the system needs to be technically adjusted to unexpected external disturbances. For instance, if there is a train accident, the traffic needs to be rerouted and actions need to be taken to repair the damage as soon as possible. This requires tight supervision in order to force the necessary changes of the technical state of the system. Specific investments are required in order to reduce uncertainty and prevent strategic behaviour, since a failure of the operational balancing will have direct consequences for the reliability of infrastructure systems.

At T_5 capacity utilization requires monitoring of the operation of the technical process. Resources need to be dedicated for the production process in order to produce the desired infrastructure services. Since this allows for some relocation between different production means, the degree of asset specificity is declining, along with the degree of

[26] Still asset specificity can be an issue at this level of analysis. Components like electricity plants represent long term investments of several decades.

[27] Corroboration is here understood as a process of securing and/ or strengthening complementarities between the components of the system.

uncertainty and opportunities for strategic behaviour. For example there may be backup capacity in electric power plants that can meet peaks of demand in exceptionally cold or exceptionally warm weather.

T_{15} denotes a point in time in which capacity allocation is possible. This long term process needs to be facilitated in order to make sure that critical transactions are supported to sustain the technical system complementarities. Since there is a long time period involved to perform the feedback loop between the expected and actual system performance, there is opportunity to deploy technological and organizational variety. Referring to the electricity sector example, new power plants can be built, or the networks extended or reinforced. However, in order to support system complementarities, this process needs to be facilitated in order to make sure that the desired outcomes are realized. Building new power plants might make it necessary to adjust network capacity. At this point, there is limited asset specificity involved. However, there is significant uncertainty with respect to the future technical and economic states of the system, which opens the door to strategic behaviour. Examples of sources of uncertainty are the future demand for energy, or future preference with respect to primary energy sources, like fossil fuel or sustainable sources like wind or solar.

T_{50} is concerned with system transformation and/ or innovation. This is associated with planning in order to meet technical and organizational needs of the system. New water facilities, such as building a new dam or creating new reservoirs, or building new gas pipelines that cross different countries are illustrative. This is the most long term process with respect to the establishment, operation, and management of control systems. Asset

specificity is submitted to major trade-offs, since many alternative investments, possibly by different and even competing operators can be considered, whereas uncertainty is high.

Combining these different values of the technical-control dimension leads to twelve possible modes of organisation that are indicated by the cells of Figure 2. These modes of organization imply a certain structure or 'logic', which is visualized by different degrees of shading, basically from the most centralized forms of coordination to decentralized and possibly competing ones.[28] The upper left (dark shaded) cell indicates the case of 'authoritative supervision', typically through hierarchical integration or similar arrangements. Since system wide critical transactions are involved and technical balancing needs to be assured on the spot, only a centralistic and prescriptive approach of organization can meet these requirements. Information systems and backup equipments for coordinating airlines are an absolute requisite to guarantee effectiveness and safety of air traffic. This can be understood as the most critical transaction. The other extreme is the bottom right cell of 'decentralized planning' of system innovations at the level of components, e.g., developing new models of locomotives. This very long term process allows for very different modes of governance as long as the result of this decentralized planning process meets the needs of infrastructure systems. In the most far reaching case this could be left to autonomous market processes, in which the 'invisible hand' would act as a selection process for the innovations to be performed. In between (lightly shaded cells) we have different forms of hybrid modes of organization, as with mutual monitoring, in which autonomous parties have to share some decision rights, and eventually, ownership

[28] There are other important aspects that influence the alignment between technical functions and modes of organization that are not considered here, e.g., the role of market structures, the impact of public policies and regulation, or the effects of radical innovations.

over some specific assets that require coordination. Allocating tradable rights of access to networks illustrate this point.

V. DIFFERENCES ACCROSS INFRASTRUCTURES AND OVER TIME

Having elaborated the concept of critical transactions and related it to different modes of organization, we now briefly indicate two important issues that could help analysing the relation and tensions between technological requirements and modes of organization, particularly in a context of reform of infrastructures. First technology in different infrastructure sectors imposes different technical control systems in order to support different degrees of technical complementarities. Second, technological innovation may have a deep impact on the way transactions can be effectively organized.

V.1 Differences across infrastructures

Different technological needs translate into specific critical transactions, and hence different modes of organization that are available for a possible restructuring of distinct sectors.

Referring to the technical-control dimension, infrastructures differ with respect to scope of control and speed of adjustment. The prevalent technology of the electricity sector constitutes an extreme case of operational balancing of the entire system. The inflow and outflow of power needs to be technically balanced within very short periods of time in order to safeguard the reliance of the system. Under the given technological conditions, there is no opportunity for time shifting of the consumption and/ or production of electricity, since there are no large-scale means of storage or demand side management. This creates very complex technological and organizational needs in order to support the

technical functioning of the system. The railway sector, in contrast, is less stringent with respect to the need for operational balancing, since there are technical opportunities to shift the operations of trains in time by allowing for certain delays. This diminishes the reliability of the service, but this does not threaten the technical dependability and safety. The technical consequences of a delay of a train service are far less serious as compared to a blackout in an electricity system. Indeed, passenger transport services can be shifted in time, electricity consumption usually not. On the other hand, the slot allocation of tracks and railway stations is far more restrictive compared to the transport and distribution of electricity. Electricity is a homogenous commodity that is transported throughout the network according to physical laws without a predetermined and dedicated pathway to be followed. Obviously, this is a very different case for the personal transport services of railways. Certain tracks and stations need to be allocated at specific programmed dates and times. As a consequence, in this sector timetabling is a very critical transaction imposing restrictions with respect to the modes of organization of capacity utilization and allocation.

There are also infrastructures in which there are only limited inherent critical transactions. Radio frequencies constitute an interesting case particularly with respect to scope of control. Operational balancing is necessary to prevent technical interference between the frequencies. In order to solve this problem, radio frequencies are exclusively allocated for different purposes like emergency services, army, aviation, shipping, broadcasting, wireless telephony, wifi, etc. Dedicated users receive a right to use the frequencies at certain times and locations. In other words, subsystems are created that function independently from each other on an exclusive basis. It is up to the users to utilize and allocate services according to the technical opportunities. In order to protect the technical integrity of the entire system of radio frequencies, certain norms and standards are

defined with respect to the various components that constitute this infrastructure. Under these circumstances operational balancing is mainly based on general framework conditions. The most sensitive issue for the use of frequencies is capacity allocation. Liberalizing this sector would involve a discussion on alternative modes of capacity allocation, for instance guided- or competitive allocation mechanisms. However, a change of the mode of organization of a specific critical transaction can have implications for the mode of organization of other critical transactions. Changing for instance the allocation mechanism of radio frequencies towards competition, might involve a necessity to strengthen supervisory authorities that monitor the operational balancing of the use of frequencies.

To sum up, our classification of modes of organization of critical transactions suggest that (1) different infrastructures imply different critical transactions; (2) different opportunities for sector restructuring depend on feasible modes of organization of critical transactions; and (3), in order to guarantee reliable system services, modes of organization in a specific infrastructure need to satisfy its critical technical functions. Any reform that intends to be effective without threatening the efficiency of the system must incorporate these criteria.

V.2. A dynamic perspective

So far we have considered the technology of infrastructures as given. However, technology changes with implications for the critical transactions thus requiring adaptation or even radical changes in modes of organization, which in turn imposes built-in flexibility in order to make a reform durable. For instance, information and communication technology increasingly offers opportunities to monitor and control network related

activities, thus lowering transaction costs and making viable market arrangements that previously involved prohibitive costs. In the electricity sector for instance, the critical transaction of operational load balancing might be shifted from the system level (i.e. high voltage network) to the subsystem (i.e. low voltage distribution network). From a technical perspective this would imply the introduction of so called intelligent distribution networks ("smart grids"). Nowadays distribution networks have no other technical function than transforming the voltage level (for instance from 380 KV to 220 V) and transporting electric power to the final customers. Future intelligent electricity distribution networks would perform important system tasks like facilitating the production of electricity in small scale distributed plants, and technically balancing the inflow and outflow of power. This would result in an electricity web, similar to the internet, with partly technical independent subsystems that are connected through a backbone in order to allow power exchanges between the different subsystems. If this system innovation is to be realized, this would have significant implications for the modes of organization of critical transactions, since new opportunities for restructuring arise, e.g. more decentralized and less tightly coordinated.

Another remarkable example is road pricing based on electronic portals and automatic debiting systems. With these new technologies, physical access control and payment is not anymore necessary. This results in a more efficient system of monitoring and control with less transaction costs, but it also involves tighter coordination with respect to information flows and controlled access.

These examples illustrate that technological change has fundamental repercussions for the nature of critical transactions and accordingly the possible modes of organization.

Restructuring infrastructures without considering the nature of its critical transactions and the according modes of organization can be doomed to result in major flaws with respect to the technical reliability of services.

VI. CONCLUSION

In this paper we addressed the interrelation between technology and institutions in infrastructures, with a view at sector reform-oriented policies. Often these institutional reforms neglect the technical dimension of infrastructures as well as the organizational requirements as possible barriers or drivers for change. This can result in inferior technical performance, in some cases even resulting in major accidents like blackouts in the electricity sector, accidents in rail transport or substantial leakages in water systems. Since infrastructures are characterized by strong technical complementarities, there is a need to sustain and harmonize technical processes that are physically interlinked through networks. Our analysis suggests that what is crucial is the capacity to align reforms in modes of organizations with critical transactions in order to preserve a coherent framework in which critical technical control functions are related to a specific set of modes of organization so as to guarantee reliable system services.

Our analysis of critical transactions also points out that different infrastructures imply different critical transactions. Hence, depending on the possible modes of organization of critical transactions, there are different opportunities for sector restructuring. We even went a step further: not only are there important differences across sectors, but there can also be significant differences in criticality among different components of a specific sector, as illustrated by our figure 2, so that modes of organization

need to be conceived as modules that must be properly aligned with each other according to the characteristics of the transactions at stake.

The successful restructuring of infrastructures strongly depends on the capacity to align technical functions and the modes of organization by properly identifying critical transactions. Reformers must also keep an eye on how technological innovation can interact with existing arrangements. Therefore, our analysis has strong implications for policy makers: critical technical functions need to be clearly identified; critical transactions should be differentiated from transactions that do not threaten the coherence of a system; and the modular aspect of organizations dealing with complex nodes of transactions should all be taken into account. Restructuring infrastructures without keeping simultaneously a grip on these dimensions is doomed to result in major failures with respect to the technical reliability of services. Policies oriented towards successful reforms of key infrastructures should urgently be revised in order to integrate these dimensions.

REFERENCES

Arthur, B. W., 1988. Competing technologies: An overview. In: Dosi, G., Freeman, C., Nelson, R. R., Silverberg, G., Soete, L. (Eds.), Technical change and economic theory. Pinter, London.

David, P. A., 1985. Clio and the economics of qwerty. American Economic Review 75, 332-337.

Dosi, G., 1982. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. Research Policy 11, 147-162.

Dutton, K., Thompson, S., Barraclough, B., 1997. The art of control engineering. Addison Wesley Longman, Harlow, UK.

Economides, N., 1996. The economics of networks. International Journal of Industrial Organization 14, 673-699.

Finger, M., Groenewegen, J. P. M., Künneke, R. W., 2005. The quest for coherence between technology and institutions in infrastructures. Journal of Network Industries 6, 227-259.

Geels, F. W., Schot, J., 2007. Typology of sociotechnical transition pathways. Research Policy 36, 399-417.

Gourvish, T., 2008. Britain's railways 1997-2005. Labour's strategic experiment. Oxford University Press, Oxford.

Hines, P., Apt, J., Talukdar, S., 2009. Large blackouts in North America: Historical trends and policy implications. Energy Policy 37, 5249-5259

Institute of Electrical and Electronics Engineers, 1990, IEEE standard computer dictionary: A compilation of IEEE standard computer glossaries, New York.

Joskow, P. L., 1997. Restructuring, competition and regulatory reform in the u.S. Electricity sector. Journal of Economic Perspectives 11, 119-138.

Joskow, P. L., 2003. Electricity sector restructuring and competition: Lessons learned. Cuadernos de Economia 40, 548-558.

Klein, P. G., 2005. The make-or-buy decision: Lessons from empirical studies. In: Ménard, C., Shirley, M. (Eds.), Handbook of new institutional economics. Springer, Dordrecht, 435-464.

Kroes, P., Franssen, M., van de Poel, I., Ottens, M., 2006. Treating socio-technical systems as engineering systems: Some conceptual problems. Systems Research and Behavioral Science 23, 803-814.

Ménard, C., 2005. A new institutional approach to organization. In: Ménard, C., Shirley, M. (Eds.), Handbook of new institutional economics. Springer, Dordrecht, 281-318.

Moteff, J., Copeland, C., Fischer, J., 2003, Critical infrastructures: What makes an infrastructure critical? Report to the Congress. Congressional Research Service.

Murmann, J. P., 2003. Knowledge and competitive advantage: The coevolution of firms, technology, and national institutions. Cambridge University Press, Cambridge.

Nelson, R. R., 1994. The co-evolution of technology, industrial structure, and supporting institutions. Industrial and Corporate Change 3, 47-63.

Nightingale, P., Brady, T., Davis, A., Hall, J., 2003. Capacity utilization revisited: Software, control, and the growth of large technical systems. Industrial and Corporate Change 12, 477-517.

North, D. C., 1990. Institutions, institutional change, and economic performance. Cambridge University Press, Cambridge (Mass).

Perez, C., 2002. Technological revolutions and financial capital: The dynamics of bubbles and golden ages. Edward Elgar, Cheltenham, UK.

Salanié, B., 2005. The economics of contracts: A primer. MIT Press, Cambridge (Mass.).
Saviotti, P. P., 1996. Technological evolution, variety, and the economy. Edward Elgar, Cheltenham.
Shelanski, H. A., Klein, P. G., 1995. Empirical research in transaction cost economics: A review and assessment. Journal of Law, Economics, & Organization 11, 335-361.
UN Millennium Project, 2005. Health, dignity, and development: What will it take?, Earthscan, London.
Van de Poel, I., 2003. The transformation of technological regimes. Research Policy 32, 49-68.
Von Tunzelmann, N., 2003. Historical coevolution of governance and technology in the industrial revolutions. Structural Change and Economic Dynamics 14, 365-384.
Williamson, O. E., 1985. The economic institutions of capitalism. The Free Press, New York, London.
Williamson, O. E., 1996. The mechanisms of governance. Oxford University Press, New York.
www.gridwise.org, 2007, Gridwise.Org.
www.smartgrids.eu, 2007, Smartgrids: European technology platform.