Free Entry in Infrastructure

David Ehrhardt and Rebecca Burdon[1]

[1] David Ehrhart and Rebecca Burdon work for London Economics, which was contracted by the World Bank to prepare this report.

Table of Contents

Tables & Figures

Page

Free entry – Executive Summary

In most countries, a single firm has a legal monopoly on supply of water, electricity and sanitation in any given area. No one else is allowed to supply such services.

Why do such rules exist? Do they make sense, or would it be better to allow anyone who wished to supply these services? These are the questions this paper answers.

What is free entry?

Free entry means letting anyone who wants to supply services do so.

Free entry is a situation in which individuals, firms or community groups who wish to supply power, water, and sanitation services are allowed to, with no (or minimal) legal restrictions. Free entry is the opposite of 'exclusivity' or 'legal monopoly' in which only one firm is allowed to supply the service.

By 'no restrictions' on entry we do not mean 'no regulation'. Infrastructure providers will always have to comply with laws on investment, company registration, planning etc. which apply in all sectors of the economy. They may also have to comply with industry specific laws, for example on technical standards.

Free entry in this paper means that anyone who complies with such laws is free to supply utility services, and that any industry-specific laws will be fair to all firms in the industry, and not biased in favour of an existing supplier.

Why is free entry banned?

Free entry is allowed in most industries ...

… but usually Governments only allow one provider of power, water and sanitation in any area.

This is supposed to ensure supply to poor and marginal areas …

For most products and services, even essential ones, competition is usually valued and encouraged. Most neighbourhoods are served by several competing food shops. Fears that incumbent airlines in the US may be crushing start-up airlines through predatory pricing have led the Department of Transport to take action to protect the entrants. There are many other examples.

In water and electricity, by contrast, Governments usually ban competition. Often, the objective is to provide service to poor and marginal groups. A single utility is mandated to provide services to an entire area, including poor and marginal areas which might not be able to pay the full cost of supply. To cover the costs of this social service, the company needs a secure base of more affluent customers, who can be charged more than the cost of supply.

Competition could undermine this.

… and to prevent wasteful duplication.

Governments also fear that having multiple providers could result in wasteful duplication of infrastructure, or poor co-ordination of investments and operations.

Water and electricity supply is often a natural monopoly. This means a single company should be the technically most efficient provider. A single company can reap economies of scale, and ensure that investments in various parts of a complex, wide-spread and expensive network are properly co-ordinated.

Having a single utility can also make it easier ensure that technical, safety, quality and environmental standards are met; regulators need only supervise a single entity. In some cases, exclusive rights are granted to promote private investment in the utility.

Should free entry be allowed?

In giving a single utility a legal monopoly, governments hope it will operate and invest efficiently, cover its costs, and provide a good quality, reliable service to the entire area it serves, including slums, low income areas, urban peripheries, etc.. Where these hopes are fulfilled, there may be little reason to tamper with success. This is especially so in countries with limited governance capacity, where managing change and competition could introduce real risks.

Often the monopoly provider does not supply poor and marginal areas.

Unfortunately, in most parts of the world these hopes are not fulfilled. Monopoly water and power utilities often operate at high cost, lack funds to invest, and provide a low quality, unreliable service. Worst of all, they frequently provide limited or no service to poor and marginal areas. When the monopoly model is not working, it is time to look at alternatives.

Looking at developing countries where alternative solutions have developed gives an idea of what could be possible. Here are three examples

Alternative providers can enter the market to fill the gap.

Orangi Pilot Project - Sanitation *- Orangi is an unplanned settlement in Karachi, Pakistan. A population of around 800,000 working class people lived in an area with medieval sanitation conditions. Diseases such as typhoid, malaria, diarrhoea, dysentery and scabies were rampant. The population waited for years for the municipal utility to provide a subsidised sewerage system, but it never came.*

Starting in 1980, a charitable group developed a low cost approach to piped sanitation. The group explained the technology to the community, and catalysed community action. Householders and

neighbourhoods funded and managed the construction of household pourflush latrines and sewerage lines. By 1993 the services extended to about half the settlement. Householders have benefited from improved health and increased property values, and continue to contribute the time and money necessary to keep the system working.

Paraguay's aguateros – water supply *– There are about 300 to 400 private individuals and firms in Paraguay supplying good quality piped water to areas not served by the public water company. These range from 'mom-and-pop' operations supplying their own neighbourhood, up to larger companies serving as many as 800 connections.*

Unlike the public company, the aguateros allow payment of connection fees on instalment, making it easier for low income consumers to connect. A UNDP-World Bank Water and Sanitation Program review concluded "Aguaterias have been operating successfully for a good ten years and quality of service is increasing as experience, competition and customer activism have increased over time".[2]

Yemen's small scale electricity providers *- In the Yemen, the public utility meets urban electricity demand, but does not supply rural towns and villages. This gap is filled by a large number of small-scale providers. These range from individual households who generate for their own use and supply a few neighbours, to larger operators supplying up to 200 households and operating sophisticated billing systems. The result is that rural household electricity use in the Yemen is remarkably high compared to other middle income countries. While there is little prospect of significantly improving the proportion of rural grid-connected households in the next two decades, small scale private sector operators have demonstrated a willingness and innovative approaches to meeting the demand.*[3]

These examples show that where a monopoly supplier does not meet people's needs, other suppliers can enter the market. These suppliers are generally financial viable, sustainable, and customer responsive. They adopt innovative technology, and expand service to poor and marginal areas. Given a chance, such small-scale enterprises can evolve into substantial providers, and generate

---

2 Solo, Tova Maria and Snell, Suzanne Water and Sanitation Services for the Urban Poor - Small Scale Providers: Profiles and Typology UNDP-World Bank Water and Sanitation Program, May 1998.

3 Republic of Yemen: Household Energy Strategy Study, Phase 1. A Preliminary Study of Northern Governorates. Report No. 126/91 Energy Sector Management Assistance Program (ESMAP) March 1991.

significant welfare gains in the communities they serve. This suggests that restricting entry into crucial sectors such as water and electricity may impose real costs, in particular on the poorest communities.

These entrants seldom duplicate investments …

Entrants tend to concentrate on unserved areas, so that in practice wasteful duplication of investment seldom occurs. Our analysis suggests that it is only when there is no existing infrastructure, and the market is likely to be highly profitable, that a race for market share can result in duplication. These conditions exist in some markets, for example in the wireless telephony market in many countries. However they are rare in power and water supply in developing countries.

… though some Government intervention to ensure interconnection is possible could boost efficiency.

Co-ordination between networks is a bigger problem. The low cost sewerage system of Orangi needs to connect to the main municipal sewer lines so that the effluent can be carried away from the neighbourhood for disposal. If an entrepreneur in one area establishes a small electricity system which uses a 60Hz cycle, and an entrepreneur in a neighbouring area establishes one running on 50Hz, it will be impractical to interconnect the two systems in the future as they expand, so future cost savings and gains from trade will be lost. These issues suggest that there may be a valuable role for Governments in setting technical and interconnection standards in some cases.

Limitations on entry may sometimes be justified for environmental reasons …

There could be times when allowing free entry could worsen environmental outcomes. For example, numerous small abstractors could overwhelm a country's environmental regulatory capacity, and over-pump an acquifer, causing saline intrusion. Whether or not this is likely to be a problem should be assessed case by case, and the probable scale of the problem compared to the environmental and health benefits which free entry can deliver by improving water and sanitation services.

… or to promote private sector investment …

… but these cases are probably rare.

It is also true that sometimes exclusivity will be needed to attract private investment. However, generally governments will be able to limit the length of time for which exclusivity is given, and should consider other options which could attract investors without prohibiting entry.

Experience in many countries, including in Argentina and Guinea in the water sectors, shows that privatising monopolies generally improves efficiency, but the privatised companies still lack an incentive to extend cross-subsidised service to low income and peripheral areas. Even with a dominant private supplier, allowing free entry could be beneficial.

Conclusion

Legalising alternative providers will allow them to expand and meet new needs.

Limits on entry will sometimes be needed, but these should be exception, not the rule.

Generally, free entry should be allowed in power, water and sanitation.

Currently, many developing countries prohibit entry into water and electricity markets. All too often, the monopoly supplier provides a poor service, and fails to supply many low income and marginal areas.

Alternatives do exist. These range from the very small scale, like water carriers and privately-operated pay toilets, up to networks supplying hundreds of customers with power, water and sanitation services. These alternatives are generally created by entrepreneurs or community groups. They are usually innovative, low cost and sustainable.

Given the valuable services such entrants can provide, it is generally counter-productive to outlaw them. Yet this precisely what legal exclusivity for a single utility does.

Our conclusion is that there should be a presumption in favour of free entry. That is Governments should generally remove legal restrictions on the establishment of new water, sanitation and electricity suppliers. Removing such restrictions will make it easier for existing operations to expand, and for new entrants to target market niches where needs are as yet unmet.

Free entry will not always be the best policy. In some cases the risk of duplication, environmental degradation, or discouraging private investment, will justify limits on entry. But these cases are likely to be rare. Governments should allow free entry, unless there is a clear reason to limit it. This would be a reversal of the current policy position in many countries.

1 Introduction

The World Bank asked London Economics to advise on the circumstances, if any, in which governments should restrict entry into infrastructure industries that are likely to be natural monopolies, such as piped water and electricity transmission and distribution. The study focuses on the water and electricity industries on the assumption that water and power are the most important infrastructure services for low income consumers in countries with poor governance. Information and examples from other network infrastructure industries are also discussed.

The study assesses the costs and benefits of restricting entry, and draws policy implications from that analysis. We do not focus on whether free entry in infrastructure is likely to occur if allowed, nor if it will be an important factor in infrastructure provision in the future. Likewise, we do not consider whether developing countries with poor governance should adopt policies to promote competition over a network, for example by creating power pools.

We begin by defining 'free entry' and then briefly outlining current policies towards entry in infrastructure industries. We show that in the majority of countries, free entry is not permitted.

The general prohibition on free entry is out of line with recommended policy for most other sectors of the economy. Generally, allowing competitive entry to a sector will promote efficiency.

In Section 3 we propose a general presumption that entry to infrastructure sectors should be freely allowed. We show why this could be expected to be beneficial, and give examples of where it has occurred.

In Section 4 we go on to consider the problems and unintended results which, potentially, could follow from allowing free entry in infrastructure. Such potential problems could include:

- inefficient duplication of facilities;
- loss of universal service;
- difficulty in attracting private sector investment;
- poor co-ordination of systems, and inefficient scheduling and routing decisions; and
- deterioration in environmental quality and safety outcomes.

In each of these areas we:

- review the problems which theory suggests could occur if free entry were allowed;
- outline circumstances where the problems may be most likely to arise, and provide practical examples of such problems;
- so far as possible we draw conclusions as to the likely prevalence and scale of the problems; and finally
- review the likely effect of prohibiting free entry, taking into account dynamic and comparative institutional issues.

Free entry cannot be considered in isolation. It interacts strongly with other complex systems, in particular:

- technical standards and the development of integrated systems;
- regulation; and
- political dynamics.

In Section 5 we consider the implications of free entry for these associated systems. In Section 6 discuss the likely significance of adopting a policy of free entry, and finally, in Section 7 we summarise our conclusions.

2 Background: current policies on entry into infrastructure

2.1 Defining 'free entry'

It is difficult to provide a single, comprehensive definition of 'free entry'. The essential idea is that free entry with respect to a particular service refers to a situation in which there are no legal barriers to supplying that service.

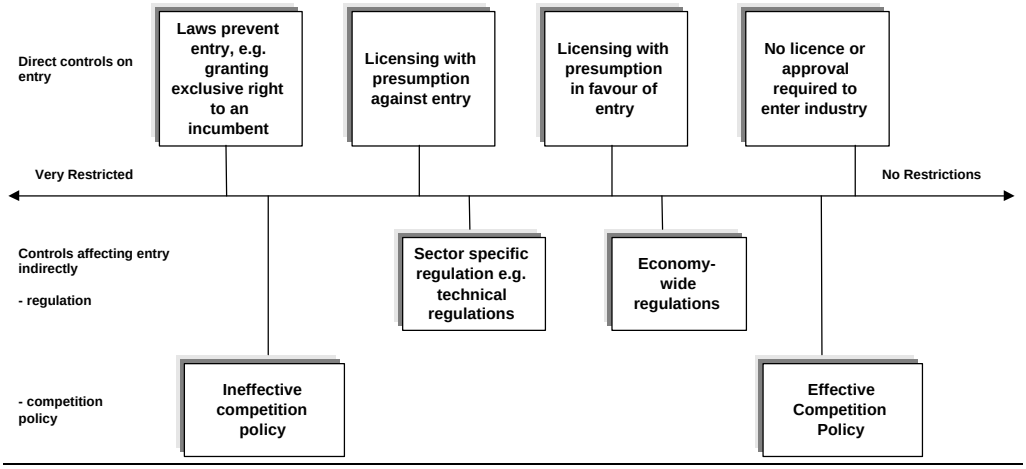
In practice, there are always some legal provisions which could be interpreted as legal barriers to entry. Even in the most liberal regime, entrants will often have to comply with general legal provisions requiring, for example, registration of companies and filing of tax returns. However minor, these could be interpreted as barriers.

Taking this into account, a possible definition of free entry would be "a situation in which entry into the industry in question is no more restricted than entry into other industries, for example manufacturing and retailing." This definition also breaks down, for two reasons. The first is that in countries in which entry into all industries is restricted (for example Cuba), it would be inaccurate to say that there was free entry in water and power simply because these were no more restricted than other industries.

The second is that there will often be industry-specific regulations which apply to water and power providers, but not to other industries. For example, water providers will often have to obtain water abstraction licences. This could be seen a barrier which does not exist in other industries. Nevertheless, if the sector-specific requirement is administered in a competitively neutral way (for example, if anyone can purchase abstraction licences in an auction), it would not infringe the definition of free entry that we have in mind.

Figure 2-1 captures some of these issues. Its shows a policy spectrum, from prohibition on entry to complete freedom to compete.

Figure 2-1: Policy spectrum



In Figure 2-1, specific restrictions on entry are shown above the line. The continuum moves from legal exclusivity for a single supplier, through progressively more liberal licensing regimes, to a point in which there are no licensing or specific controls on entry at all.

Below the line the general regulatory environment is shown. The positioning of the various factors is illustrative only. On the whole, the more intrusive the sector specific technical regulations, the greater the difficulty of entry. Economy wide regulation can also limit entry. Generally the effect will be insignificant, but in some cases economy wide regulation will matter a great deal. High tariffs on imported small generators, for example, could seriously impede small scale entry in the electricity market. Limitations on foreign investment may block other types of entry.

Competition law is also important. Generally, effective competition laws will assist entry by allowing entrants access to bottleneck facilities, and preventing incumbents from crushing them through predatory pricing or other tactics. However there will also be times when competition law deters entry. For example, an entrant may refuse to invest in a greenfields site if it fears it will then be required to allow competitors access to its facilities on regulated terms. Sometimes market sharing agreements would be necessary to reduce risk sufficiently to encourage investment. Competition law may prohibit such agreements, deterring investment.

Finally, there is the problem of finding a definition of 'free entry' which encompasses the experiences of countries in which entry has occurred. This is difficult because often the law forbids entry in many sectors, and yet, since the law is not enforced, entry occurs.

Therefore, in this report, when we use the term 'free entry' in a theoretical or prospective sense (for example in recommending that a country allow free entry, or discussing the problems which could result if free entry is allowed) we mean a situation in which:

- there are no laws preventing entry;
- no licensing requirement (or a presumption that licences will be issued); and
- the general regulatory and legal conditions do not make entry more difficult than it would be in other industries in a typical liberalised economy.

When talking about actual examples of free entry, we use a broader definition. It includes entry which occurs even when theoretically it is illegal, and situations in which competing licences or authorisations are in fact issued, even when there was not a clear presumption that they would be. This broader definition is necessary because, as Table 2-1 shows, there are very few examples of 'pure' free entry regimes. Nevertheless, much can be learned from the cases in which entry has occurred, even when it is not fully legally sanctioned.

The specifics of licensing or regulations are not discussed in any detail as this is not the focus of the paper. In Section 5 however, we explore the implications of free entry for regulations to ensure technical compatibility, interconnectability and broader economic regulation of the sector.

2.2 Current policies on entry

Many, perhaps most, countries do not allow free entry in the provision of infrastructure facilities in the water and power sectors. Free entry is, however, increasingly permitted in other infrastructure industries, in particular in telecommunications.

Table 2-1 provides examples of where various countries and sectors lie in the range from prohibition to free entry.

Table 2-1: Spectrum of entry policies: examples

Policy / Sector	Prohibition or severe restrictions on entry*	Licensing with real prospect that entry will be permitted	De facto free entry**	Free Entry
Water	Buenos Aires, Argentina – private operator has an exclusive concession contract Santiago, Chile – public sector company with exclusive franchise Trinidad and Tobago - publicly owned, privately managed company	England and Wales (current) - 'inset appointments' allowed by regulator London (eighteenth and nineteenth centuries) - the capital's water supply was created by private companies who competed with each other where networks overlapped[i] Kenya, Nairobi – licensing of competing water kiosks[ii] Senegal, Dakar – private standpipe operators (fontaniers)[iii]	Mali - mining investor installed own water pipeline at cost of US$25 million[iv] Nigeria, Onitsha – thriving private water trucking industry[v] Paraguay (many private supply companies – 'aguaterias')[vi] Philippines (various towns) – UV disinfection based water vending stations operating[vii]	Bangkok, Thailand – household reselling of municipal water permitted[viii] Jakarta, Indonesia – household reselling of municipal water permitted[ix] Guatemala City, Guatemala – water trucking, with water supplied from private wells[x] Senegal, Dakar - water carriers and water carters[xi]
Sanitation	Buenos Aires, Argentina	Benin, Cotonou – Private septage collection by tanker, and treatment.[xii] England and Wales (current) - 'inset appointments' allowed by	Malang, East Java, Indonesia – small bore sanitation and treatment system established by private entrepreneur Orangi Pilot Project, Karachi,	

Table 2-1: Spectrum of entry policies: examples

Policy / Sector	Prohibition or severe restrictions on entry*	Licensing with real prospect that entry will be permitted	De facto free entry**	Free Entry
		regulator Jamaica – private collection of septage by tanker. Private piped treatment provided by housing developers and resorts[xiii] Nigeria, Agbara Industrial Estate has a private company treating waste for other companies on the estate[xiv] Philippines, Dagupan City – a public toilet is successfully run as a business by a private entrepreneur	Pakistan – community group catalysed provision of low cost piped sanitation Lahore, Pakistan – based on Orangi Pilot Project Trinidad – majority of sewerage provision is through small private systems installed by housing developers, septic tanks, and private trucking of septage[xv]	
Electricity	Argentina – transmission and distribution Nigeria (most places)[xvi]	UK – Power generation, electricity retailing Nigeria (Maroko, Victoria Island – entry was permitted in unserved area)[xvii]	Mali – mining investor installed own electricity system[xviii] Yemen- a large number of small operators supply nearly all rural electricity.[xix] Laos- a significant proportion of electricity in rural areas is	Argentina – power generation and marketing (retail supply) Australia – generation– no barriers to entry in Victoria, New South Wales, Queensland or South Australia[xxi] Australia – retail supply – barriers

Table 2-1: Spectrum of entry policies: examples

Policy / Sector	Prohibition or severe restrictions on entry*	Licensing with real prospect that entry will be permitted	De facto free entry**	Free Entry
			supplied by small generators operated by households communities or private entrepreneurs[xx]	to entry being removed in most states[xxii] New Zealand[xxiii]
Other	Jamaica – Telecoms (Cable and Wireless has 25 year exclusive licence over most telephone services, with a further right of renewal)[xxiv]	Australia – State of Victoria – gas retailing. Licensed entry will be permitted[xxv] Australia telecommunications[xxvi] Nigeria – postal services (entry by DHL permitted)[xxvii] UK – gas retailing, telecommunications US telecommunications (many segments) US railways in the nineteenth century[xxviii]	International call back services in many countries Internet telephony in many countries	Argentina – gas production and marketing (retail supply) Australia – gas production in Victoria, New South Wales[xxix] New Zealand – gas, Telecoms, postal services[xxx] Chile – gas transmission Global communication satellite systems[xxxi]

*Given the number of countries in this category, we only list those referred to in the text

** By de facto free entry we mean that entry may not be legally permitted, but it occurs. We include in this column cases where we know small scale entry has occurred, but we are not sure of the legal framework. Thus it is possible that some of the de facto free entry examples are in fact cases of legal free entry.

Note: all references footnoted in this table are contained in Annex 4 at the end of the paper

As far as we are aware, most of the countries which legally sanction free entry in their formal water and power sectors are developed nations. In most cases a licence is required and free entry is restricted to certain areas. For example:

- most Australian States permit free entry in electricity generation (although licences are needed to join power pools in most States), and are liberalising entry in electricity retailing;

- the UK allows entry in: gas (retail and network extensions); electricity (retail and new connections but not network extensions); and water distribution through 'inset appointments'. In all of these cases a licence is required, but the main requirement is that basic minimum technical standards are met.

- New Zealand allows entry in electricity, gas and telecommunications. Entrants are generally required to register with the Government, and comply with competitively neutral regulations such as information disclosure requirements.

Free entry was important in the development of infrastructure during industrialisation in the UK and USA. For example, London's water supply system was constructed by a number of private companies most of which had non-exclusive statutory authorisations from the Government, and competed with each other in a variety of ways.[4] Similarly in the US, early railroad investment often required legal authorisation, but these authorisations could be non-exclusive, and railroads competed with each other.[5]

Examples of less developed countries that allow free entry in formal and large scale infrastructure provision include:

- Argentina, which allows competitive entry in power generation and in electricity and gas marketing (retail supply). The transmission and distribution networks are the subject of concessions, but must provide access to competing producers and retailers; and

4 Tynan and Cowen (1998).

5 Chernow, Ron (1997) p.22-25.; and Wasserstein, Bruce (1998) *Big Deal: The Battle for control of America's Leading Corporations* Warner Books Inc. p.29-35

- Chile allows free entry into the gas transmission pipeline business, which resulted in competition between developers – the so-called 'pipeline war';[6]

Many countries permit legal or de facto free entry into small scale and informal provision of infrastructure-related services. A UNDP-World Bank study highlights the reasons for this, and its importance;

> *"Municipally operated public standpipes, public toilet facilities and public baths are often poorly maintained when they are not out of service entirely. … Thus it is not surprising that the majority of marginal urban residents buy water and sanitation services from small private providers who do what public providers cannot or will not." (*Solo & Snell 1998 p.1)

Services in which legal or de facto free entry is common include:

- self supply of water through wells and rainwater tanks;
- water-carrying;
- water trucking;
- night soil collection;
- cesspool emptying;
- self-supply of back-up electricity generation;
- messenger and courier services; and
- mini-bus services.

It is clear however that in the vast majority of cases free entry into significant infrastructure facilities is not permitted.

6 Warwick Smith, World Bank – personal communication

3 Presumption in favour of free entry

A natural monopoly can be defined as an industry in which it is most efficient to have a single supplier. That is, an industry is naturally monopolistic if for all industry outputs, the cost function is subadditive (i.e. the cost of producing the total output is less than the sum of the costs of producing the parts).[7] However, the fact that an industry is characterised as a natural monopoly is not in itself a rationale for prohibiting entry to the industry.

If it is most efficient for a single operator to supply a service, then, in most cases any operator entering will have higher costs than the existing operator.[8] Therefore, the entrant will not be able to compete. The single operator will remain in business, and the entrant will fail. Free entry will result in the efficient market structure of a single provider.

On the other hand, if the entrant succeeds in out-competing the incumbent, it will generally show that the incumbent was not the most efficient provider. Again, free entry will produce the efficient result.

It could be argued that allowing competition in naturally monopolistic industries is inherently inefficient because for the period in which the competitor tries to enter, it will be supplying at higher than efficient cost, creating a net loss to society. This problem is discussed in Section 4.1. However in general, these losses will be suffered by the business attempting entry. Business people do not usually make investments which they expect will result in losses.

The prima facie conclusion is that free entry is unlikely to reduce efficiency or welfare. In sectors which are true natural monopolies, free entry will usually have no effect. Free entry may increase efficiency, by allowing competition in areas that were considered to be natural monopolies, but in fact are not. The competition that has resulted in industries previously assumed to be natural monopolies is a strong argument in favour of permitting free entry.

This prima facie result is consistent with the general policy prescription in almost every other area of the economy; that competitive pressure increases efficiency by:

- reducing monopoly inefficiencies;
- increasing incentives to innovate; and

---

7 Where. $\Sigma y_i = y$, $C(y) < \Sigma C(y_i)$

8 The case in which part of the market can be supplied at a lower average cost than the whole market, even with the existence of subadditivity, is discussed in Section 4.1

- reducing reliance on regulation and government intervention and thereby decreasing risks of costly regulatory failure.

The following sections develop each of these points.

3.1 Reducing monopoly inefficiencies

The impact of competitive pressure on monopoly inefficiencies is developed in the theory of contestable markets. It emphasises that in a perfectly contestable market, a free entry policy will oblige the incumbent monopolist to operate efficiently, such that it earns zero economic profits and consumer welfare is maximised.[9] This can hold even in naturally monopolistic industries.

Efficient behaviour results from the monopolist's strong incentive to deter entry. Deterring entry requires the monopolist to adopt prices and production levels such that there are no opportunities for entry that appear profitable to entrants.

The theory of contestable markets shows that for free entry to motivate the monopolist to operate efficiently, entry and exit must be cost-less. Exit is cost-less where there are no sunk costs associated with entry (fixed costs do not constitute a cost of exit and are therefore not a barrier to entry).[10] These conditions do not hold in the case of infrastructure industries and in particular in water and electricity. The existence of sunk costs means that a free entry policy will not in itself result in the incumbent monopoly operating efficiently.

However allowing entry puts an upper bound on monopoly inefficiencies (provided the costs of the inefficiencies are borne by consumers and not just by governments). Examples include:

- the threat of major customers developing alternative water supply systems through desalination, to remove their reliance on an incumbent's network supply. For example, in Trinidad and Tobago, major industrial customers responded to the poor quality (unreliability) of the government-owned utility's water

---

9 Conditions for a perfectly contestable market are that: entry is allowed and can be reversed without cost; potential entrants serve the same demand and use the same productive techniques as those of the incumbent firm; and potential entrants evaluate the profitability of entry at the incumbent's pre-entry prices.

10 Baumol, Panzar and Willig (1988). Note: Sunk costs are costs that are incurred even if all production ceases. They are an irrecoverable cost of entry. Fixed costs are costs that can not be reduced even in the long run by decreases in output, but they can be eliminated by stopping production completely.

supply by investigating self supply through desalination. Knowledge of this "investigation" provided strong incentives for the utility to respond to ensure adequate water supply for these customers;

- the ability for of tourist areas to develop private sewerage systems rather than relying on the monopolists' integrated sewerage systems helps put an upper limit on monopoly inefficiencies. This is observed in Jamaica and elsewhere in the Caribbean; and

- in Nigeria, state water and power utilities are highly unreliable. As a result, the vast majority of manufacturing firms provide their own infrastructure facilities, such as generators and wells. Even though this costs far more than relying on the state monopolies, it provides the manufacturers with the reliability they need. The effect of this competition on the state utilities is not clear, but it is clear the free entry in the form of self supply increases efficiency overall in this situation.[11]

3.2 Incentives to innovate and use appropriate technology

The available empirical evidence supports the view that competitive pressures provide a better incentive for innovation than an uncontested position of market power.

There are a number of theoretical reasons why this may be so. For example, Tirole shows that the monopolist's incentive to innovate is less than the incentive for a competitive firm to innovate. However, where a monopoly is threatened by competition this result is reversed because, since competition reduces profits, the monopolist's incentive to remain a monopoly is greater than the entrant's incentive to enter.[12] However the theory is not unambiguous. Schumpter argued that large and profitable corporations are more likely to advance innovation due to reasons such as superior access to capital, and ability to pool risks.[13]

Competition may also promote the use of appropriate technology. At times the technology and approach of the monopolist incumbent may not meet the

---

11 Lee & Anas (1992)

12 Tirole, J The Theory of Industrial Organisation pp 389-394

13 Schumpter, J A (1942), Capitalism, Socialism and Democracy. New York: Harper

needs of the local population. Thus it will often be desirable to allow entry to enable innovative alternative sources of supply to be put in place. For example:

- alternative technology and solutions that reflect local physical and economic conditions have been used by aguateros (owners of individually owned piped water systems) to develop water supply systems in Paraguay. Since 1980 an increasing number of aguateros have established piped systems to supply water to areas not supplied by the public sector providers (SENASA or CORPOSANA). There are around 300-400 aguateros, the largest of which have around 1,200 connections;[14]

- the Orangi Pilot Project brought sanitation to an area of Karachi, Pakistan. Technological innovation was a key factor in the success of this project. Where the incumbent utility concentrated on conventional solutions, the project organisers developed simplified design and construction methods. The community could carry out much of the construction itself, and costs were cut to a quarter of those of a conventional system;[15] and

- in Laos, where electricity is a leading export, rural electrification is currently very low compared with other Asian countries, despite evidence of significant demand. The geography and patterns of rural settlement mean that conventional grid based extensions are difficult and not viable. Innovative off-grid systems such as diesel, microhydro and solar generation that can be supplied by private entrepreneurs are more appropriate. At present small private off-grid systems provide some such systems, and should be encouraged to expand.[16]

Technical innovations are often deployed by entrants, rather than incumbents. For example, a widely acclaimed small scale ultra-violet water purification unit developed in the US in 1996 is being deployed by small scale

---

14 Drangaert, Malgarejo, Kemper, Bakalian (1997)

15 Solo & Snell (1998)

16 ESMAP 1997 *Institutional Development for OFFGRID Electrification in Laos (Draft).* UNDP and World Bank Energy Sector Management Assistance Programme. December 1997.

entrants in the Philippines, South Africa and elsewhere, not by large providers.[17]

3.3 Regulatory and government failure

Well tailored, effective interventions and regulation can in principle lead to optimal outcomes. However restricting entry, tailoring interventions and removing any competitive threat places high demands on government or the regulator to:

- ensure customers' needs are met;
- limit monopoly inefficiencies; and
- determine when entry should be allowed to occur.

The risk of regulatory failure is high.[18] The costs of regulatory failure are high even in developed countries such as the US and the UK, where regulatory ability is high (see Section 3.4 for empirical examples). The risks of costly regulatory failure are likely to be higher in less developed countries. Therefore, while acknowledging the benefits of well designed and appropriate regulation, in this section we concentrate on the problems of regulatory failure in countries with low governance capacity. We outline sources of regulatory failure, experiences from developing countries and the policy implications.

3.3.1 Sources of regulatory failure

The demand for regulation comes from the fact that utilities are monopolies. Customers have no choice but to purchase from them. This allows private utilities to charge well above cost, which may be inefficient and is certainly unpopular. Government utilities are usually not profit motivated, but they can abuse their monopoly position in other ways. For example, managers

---

17 The system was developed by Dr Gadgil of the University of California at Berkeley and won Discover magazine's 1996 Award for Technological Innovation in the Environment Category, and Popular Science magazine's 1996 Best of What's New Award. Solo & Snell (1998)

18 We refer to a 'regulator', but we use the term broadly to encompass any agency of Government that tries to make a water, sanitation or power provider serve the public interest. In other words it includes truly independent regulators of private utilities such as ETOSS the water regulator in Buenos Aires. It also includes other bodies which are charged with regulating publicly owned utilities, such as Ministers, and Boards of Directors – where these are given a public interest role not a purely commercial objective (as do, for example, the Boards of State Owned Enterprises in Australia and New Zealand, which typically have a largely commercial function and are regulated by other Government agencies).

may choose to have an easy life, rather than working hard to improve service and increase efficiency.

Regulators fail to make utilities act in the public interest for a number of reasons. These include:

- lack of information;
- lack of ability; and
- lack of incentives.

To regulate effectively requires information. Regulators need to know the actual level of a company's costs, and whether these costs could be reduced. They need to know the true extent and reliability of supply, and whether these could be increased through better use of the utility's existing resources.

The fundamental problem is information asymmetry; regulators rely for their information on managers, but managers have the incentive and ability to distort and manipulate the information. While regulators may be able to compel managers to reveal information on current performance, it is near impossible for regulators to tell by how much performance could be improved.

Regulating a utility company is a difficult task, which can require a level of expertise similar to that required to manage the utility. But technical, economic, financial and legal skills are in short supply in most developing countries. Governments and regulators can seldom pay enough to attract skilled people in this area. While Ministers or Board members are sometimes highly able, they generally have many demands on their time. As a result, regulatory analysis and decision-making is often poor.

The final problem is that regulators may not have incentives to make utilities further the public interest. These perverse incentives can take numerous forms. At times regulators may just not work hard enough to achieve all possible gains.

More perniciously, regulators may seek to use their powers for personal or political ends. A Minister may refuse to allow tariff increases because this increases his popularity in the short term, even though it deprives the utility of funds to expand to cover new areas. A Board member may demand kick-backs on major contracts, even through this increases costs to the final consumer.[19] Even nominally 'independent' regulators can have their own agendas and interests to pursue.

---

[19] McCarthy, Stephen (1994) Africa: The Challenge of Transformation, St Martin's Press, New York, New York; cited in Kerf, Michel and Smith, Warwick (1996) Privatising

A study of 58 countries found that, across the economy as a whole,

> *"… corruption and bureaucratic efficiency indices are significantly and robustly negatively correlated with investment even controlling for other determinants of investment, including the political stability index"* and that *"the bureaucratic efficiency index is significantly and robustly associated with low growth … there is evidence that institutional inefficiency causes low growth"* (Mauro, P (1995), *Corruption and Growth* Quarterly Journal of Economics, August 1995 p.681-712, p.695)

The study also found a strong correlation between the level of red tape and the level of corruption. It has been suggested that bureaucrats intentionally introduce new regulations and red tape, in order to be able to extract more bribes by threatening to deny permits.[20]

Countries with the least administrative capacity failure will typically face the greatest risks of costly regulatory failure.

The costs can be extremely high. For example a study of Nigerian manufacturing,[21] found that firms spent well in excess of 10% of their total machinery and equipment budget on their own generation, water supply and other infrastructure provision. This expenditure would have been unnecessary, had the monopoly infrastructure providers offered a reasonable service. The authors noted that the public monopoly providers "have large amounts of capital already in place but fail to deliver their service at the level required to meet demand". They concluded that,

> *"The causes of [public monopoly] infrastructural failure may be grouped into two kinds. The first is relatively well understood and relates to shortcomings of the technology used by the public sector, including problems in the day-to-day management, and operation and maintenance of the facilities. The second is more complex in nature and less well controlled, and relates to general problems with administration, bureaucracy, planning, metering, billing for services delivered, revenue collection, personnel training in the public sector and lack of appropriate incentives for management and personnel in part because of civil service pay ceilings. This second set of ex-inefficiencies* [sic] *has remained the key problem over the years …"* (Lee & Anas (1992) p.1072-3)

Africa's Infrastructure: Promise and Challenge, IBRD World Bank, Washington Box 1-2.

20 Krueger, Anne (1993) *Virtuous and Vicious Circles in Economic Development* Papers and proceedings of the American Economics Association LXXXIII (1993), p.351-56; and De Soto, Hernando (1989) The Other Path New York, New York, Harper and Row; cited in Mauro (1995)

21 Lee & Anas (1992)

The above quote cites institutional problems at both the management and regulatory level as the key reason for failure of monopoly public infrastructure provision. In short, monopoly providers of utilities will often have poor incentives to serve consumers. Regulation and Government control are used to try to make utilities serve consumers. If this worked perfectly, there would be no need to allow free entry. But in practice regulators and governments often fail. The result is the chronic underperformance by monopoly utilities observed in most developing countries today.

3.3.2 Simple rules are often best

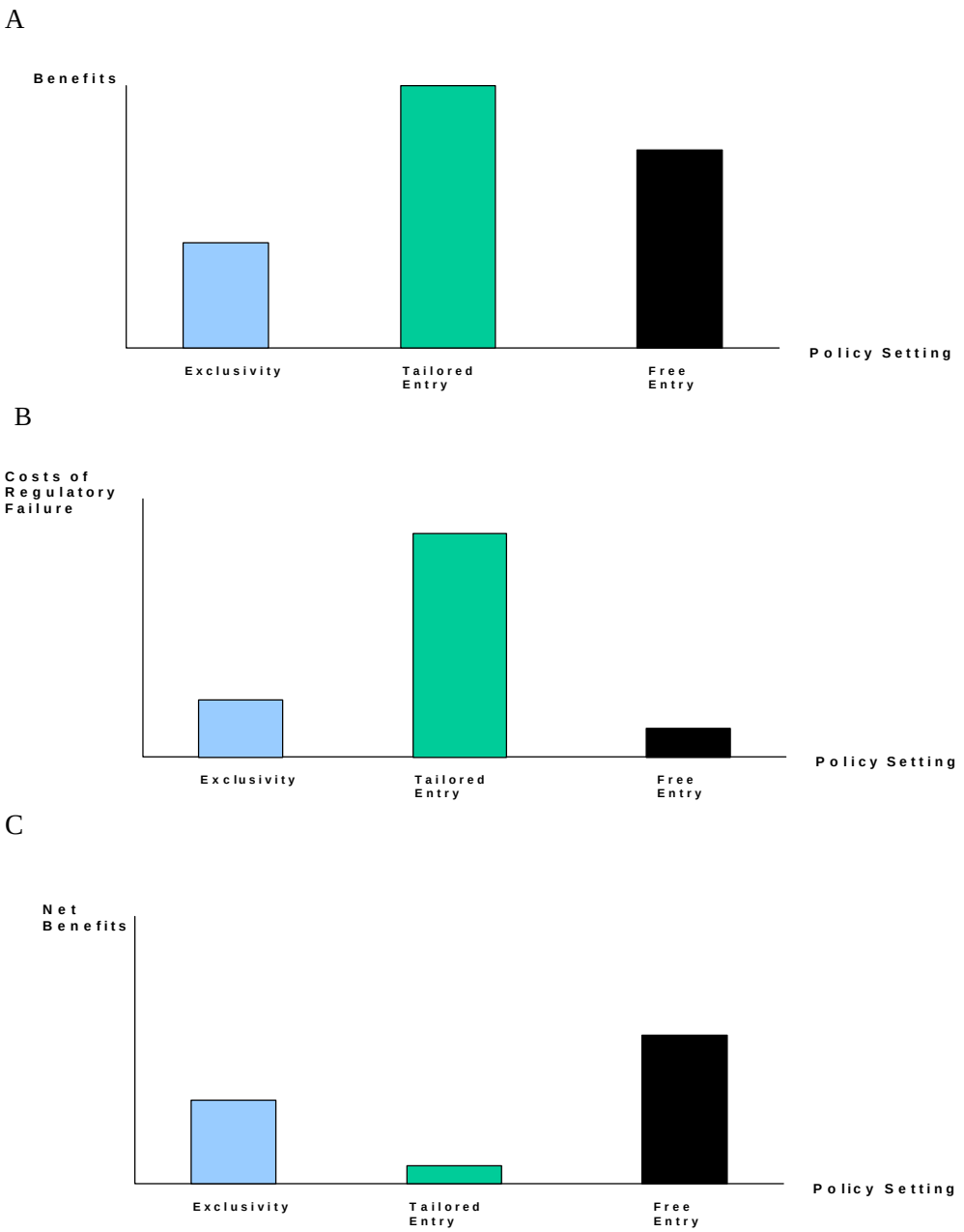
Consideration of the costs of Government and regulatory failure also suggests that, especially for countries with low governance capacity, a simple rule allowing entry in all circumstance will be optimal even if the Government knows that there are some cases in which it would be optimal to limit entry.[22]

The argument goes as follows. Despite the general presumption in favour of free entry established in this section, there are times when efficiency could be increased by limiting entry. These cases are discussed in Section 4. This is illustrated in the top diagram in Figure 3-1. The figure illustrates (in a purely subjective way) efficiency outcome from three possible policy settings: exclusivity, pure free entry, or 'tailored entry', in which the Government makes case by case decisions on the level and type of entry to allow.

Given the presumption in favour of free entry we argue that of the two extreme positions, pure free entry is likely to be more beneficial than exclusivity. Nevertheless, if we ignore regulatory and government failure, tailored entry would be better still. The Government would allow free entry where this increased efficiency, thus reaping all the gains that entry has to offer, but limit and control entry in those cases in which it could be harmful.

[22] Milgrom and Roberts (1992)

Figure 3-1: Benefits of free entry as administrative capacity varies



Once regulatory and Government failure is taken into account, the picture changes. Under tailored entry, the regulator requires a high level of skill and information in order to make effective interventions. Figure 3-1B shows the possible cost of complex analysis and discretionary decisions tailored to each situation. As governments move away from simple rules of universal application, costs rise as a result of:

- increased costs of analysis and decision making within Government, as each situation must be analysed afresh;
- increased costs on the private sector, as a result of delays and uncertainty; and
- increased losses from lobbying, corruption and other rent-seeking behaviour, as people try to manipulate the decision-making process to their own ends.

Figure 3-1B illustrates that the costs of regulatory failure are higher for discretionary systems of tailored entry than for either of the extreme positions, which involve simple rules.

We also show the costs of regulatory failure being higher under the exclusive situation than under free entry. The reason for this is that free entry allows people an escape route from the worst effects of monopoly provision, and thus mitigates the costs of failure. An additional reason is that exclusivity increases the rents available from the monopoly provider, and thus worsens harmful rent-seeking behaviour. This point is developed in Section 5.3.1.

Figure 3-1C shows the net benefits of each policy position, based on the subjective value assigned in the top and middle diagrams. The high costs of regulatory failure under tailored entry outweigh possible gains from establishing an optimal policy. Either of the extreme positions would be a better option. Of the two extremes, free entry is likely to have better efficiency effects, and lower costs of regulatory failure. On these assumptions, a simple rule in favour of free entry will be the best policy, even if free entry itself could be harmful in some cases.

Box 3-1 provides an example of a successful reform process which relied on simple rules to reduce the costs of discretionary decision-making.

---

[23] Krueger, Anne (1993) *Virtuous and Vicious Circles in Economic Development* Papers and proceedings of the American Economics Association LXXXIII (1993), P.351-56; and De Soto, Hernando (1989) The Other Path New York, New York, Harper and Row; cited in Mauro (1995)

Box 3-1: The Value of Simple Rules in New Zealand's Reforms

It could be argued that the New Zealand Government adopted the approach described above in that country's extensive economic reform programme (which commenced in 1984). The reforms relied on relatively few simple rules of general application, such as:

- allowing entry wherever possible;
- avoiding industry specific regulation; and
- relying on general competition law to govern interconnection and related issues.

These rules were undoubtedly sub-optimal in some instances.

However, the approach provided reasonable clarity, certainty and consistency, while economising on policy-making capacity. This last point is especially important for small countries undertaking many reforms at the same time, as New Zealand was.

3.4 Evidence from similar industries

The wisdom of allowing and promoting free entry has been demonstrated in many industries that exhibit economies of scale and scope, and involve network effects and scheduling problems. For example:

- following US airline deregulation in 1978, fares have fallen dramatically (an estimated 29% reduction in real prices ten years after deregulation). Other key service quality indicators such as safety records and number of interline changes required have improved since deregulation. Some estimates show the annual value of consumer benefits due to deregulation to be around $19 billion a year;[24]
- the deregulation of the US trucking industry in 1980 brought about immediate and substantial falls in rates (reduction in real fares between 1977-1987 estimated at 28% for truck load fares and 58% for less than truck load fares). Efficiency gains have been significant. The annual decline in real operating

[24] Crandall and Ellig. (1997). Economic Deregulation and Customer Choice: Lessons for the Electric Industry. The Centre of Market Processes Inc.

costs between 1987-93 are estimated to be around 9% a year;[25] and

- in the New Zealand telecommunications sector, privatisation, the removal of the incumbent's statutory monopoly, and the entry of a new carrier into the domestic and international market has brought about significant reductions in prices to both residential and business customers, and a dramatic improvement in service quality. The growth in productivity of the incumbent is estimated to have resulted in annual average cost reductions of 5.6% in the 7 years following privatisation.[26]

The benefits of competition in an industry are generally considered to be so great that private businesses with a dominant position in a market frequently have their freedom to use that position against competitors constrained by law. Good examples of such laws include the Fair Competition Act in Jamaica, the Trade Practices Act in Australia, and the competition provisions of the Treaty of Rome in the European Union.

In considering free entry, it is worth bearing in mind the sustained efforts which public authorities in a number of countries have made to facilitate competitive entry in industries with similar characteristics to infrastructure industries. High profile examples include the:

- the break-up of British Gas ten years after privatisation, to facilitate competitive entry in gas retailing;

- current US Department of Transport action designed to protect start-up airlines from being crushed by existing carriers; and

- the current anti-trust actions against Microsoft, which are intended to facilitate entry and innovation in the browser market, now a vital piece of communication infrastructure.

That in many industries with similar characteristics to infrastructure industries, the typical public policy position is to facilitate entry, raises serious questions about why in infrastructure sectors the usual policy position is to restrict or even prohibit entry.

---

25 Crandall and Ellig. (1997).

26 de Boles, David and Evans, Lewis (1995) *The economic efficiency of telecommunications in a deregulated market: The case of New Zealand* Graduate School of Business Government and Management Wellington University Working Paper 8/95

3.5 Conclusion – presumption in favour of free entry

In conclusion free entry allows customers some escape from the inefficiencies of monopoly providers. Many people around the world rely on small scale alternative providers for basic infrastructure services (even when they are technically illegally) because the dominant utility does not provide an adequate service. The possibility of entry can also provide some pressure on dominant providers to curb inefficiencies and increase innovation.

The notion that regulation can force monopolies to serve customers so well that the free entry would not be needed is flawed. Regulators and governments suffer from lack of information and ability, and may have perverse incentives. Regulation may be necessary, but it is not sufficient. Allowing free entry provides an escape route from regulatory failure.

The costs of regulatory failure and discretionary government decision-making are high. This argues for simple, clear-cut rules. The best rule will often be to allow free entry. Even where free entry could be harmful in some cases, a rule allowing free entry will often be better than trying to choose case-by-case whether entry should be permitted.

4 Potential problems associated with free entry

The literature postulates a number of potential problems that could arise if free entry is permitted. In this section we review those possible problems under the following headings:

- inefficient duplication of facilities;
- universal service issues and cross subsidies;
- exclusivity and involving the private sector;
- timetabling or routing problems; and
- effect on compliance with safety and environmental standards.

In each area we outline the theoretical problem. We describe the conditions in which this problem is likely to arise, and where possible, we provide empirical examples of the problem. We then discuss the likely costs of seeking to resolve the problem through restrictions on free entry. Where possible we also outline measures which could mitigate the problem without unduly restricting free entry, and assess the effectiveness of these measures.

4.1 Inefficient entry and inefficient duplication

In a naturally monopolistic industry, if free entry results in inefficient entry or in inefficient duplication of facilities, it could result in a net cost to society. This could occur as a result of a:

- strong economies of scale or scope across a subset of the market, which we refer to as the "*1/3, 2/3 problem*";
- a race for market position, which we refer to as the "*game of chicken problem*"; or
- two providers concentrating on a single location in the market and ignoring other parts, which we refer to as the "*two ice-cream sellers*" problem.

The following sections consider each problem in turn.

4.1.1 Inefficient entry – "the 1/3, 2/3 problem"

In most cases a naturally monopolistic market will be served by one firm, even in the absence of entry restrictions. However where there are strong product specific economies of scale or scope in a sub-set of the monopolist's outputs, a competitor could profitably supply that sub-set of the market. A descriptive illustration of this case is presented in Box 4-1, which also illustrates why we call this the "1/3, 2/3 problem".

Box 4-1: Inefficient entry

Three towns desire water supplies in given quantities. The cost of supplying any one of these towns by itself is $300, of supplying two of them via one facility is $400, and of all three towns jointly is $660. From society's perspective it is cheaper for any one company to supply all three towns, since $660 is less than the cost of supply by three separate plants ($900), or of supply by any two towns by one plant with the other town supplied by a second plant ($700).

If however, a monopolist supplies all three towns at a price per town of ($660/3=$220), an entrant could supply two of the towns more cheaply (at a cost of $400/2=$200 each).

Source: Baumol, Panzar and Willig 1982, reference to Faulhaber 1975

Box 4-1 shows that free entry could result in a situation in which there is a natural monopoly (in terms of subadditivity of production costs), but in which monopoly supply may be difficult to preserve in the free market place. This could provide a rationale for restricting entry to natural monopolies.

This problem is most acute when the conditions for perfect contestability hold (see Section 3.1). As one moves further away from these conditions, the problem diminishes. The greater the level of sunk costs, and barriers to entry, and the greater the incumbent's ability to adjust prices in response to the threat of entry, and the less likely this problem is to exist.

One would not expect to encounter this problem often in the real world of electricity and water supply. Generally, there will already be an incumbent supplying the three towns. Water and electricity supply have high sunk costs. The incumbent will generally be able to drop prices low enough to make entry uneconomic, while still covering variable and fixed costs, and earning a return on non-sunk assets. Knowing the incumbent can do this, potential entrants will not invest.

Where the sunk costs and lead time associated with entry are significant and competition law does not prohibit it, the incumbent could also price-discriminate to deter entry. The incumbent could deter entry by dropping its

price to the two towns where entry is threatened to $199. It could then charge the third town $262. In this way it could cover its full costs while offering each town a better deal than their next best alternative.

Of course the entrant could also change its entry strategy in response, and target the town which is now charged most. Whether changing the pricing policy will work depends on the response times and the commitment of the entrant and the incumbent. Where an incumbent is in place, and the entrant has yet to invest, the advantage should lie with the incumbent. The entrant will have a long lead time between starting to sink costs, and starting to supply. Once the entrant has started to sink costs, but before the system is complete, the incumbent could change its pricing policy by reducing prices to the town the entrant is targeting. The entrant would then lose the costs it had sunk. Seeing this threat, the entrant would be deterred from investing.

An entrant could get around this problem by signing a long term contract with the 1/3 customer before sinking costs. Then, even if the incumbent dropped prices, the customer would be bound to the entrant. However, in this case the commercial logic would be for the parties to reach an agreement which does not involve any investment by the entrant. For example, the incumbent should be able to supply to the 1/3 customer on the same terms as provided by its long term contract with the entrant, pay the entrant the profits the entrant would have made from its investment, and still come out better off than if investment by the entrant proceeded. However this might not work if several entrants threatened – it would not be economic for the incumbent to pay off each one.

Long term contracts can also be used to solve the problem, either by the incumbent, or when there are no suppliers yet in place and one firm wishes to capture the entire market. Any firm which can sign up all three towns on long term contracts will then be protected from entry. In practice, infrastructure investors generally ensure they have long term contracts in place before sinking investments. For example, the aguateros of Paraguay sign five year contracts with their customers before commencing construction.[27] Five years is the maximum length permitted by law, suggesting that contracts would be longer were this permitted.

In conclusion, where there are incumbents, sunk costs, and the ability to conclude long term contracts, such as in water and electricity, inefficient entry as a result of the "*1/3, 2/3 problem*" is unlikely, and does not provide a rationale for restricting entry. It is possible that it could be a concern in industries such as transportation and solid waste collection where hit-and-run entry is possible, and long term contracts are difficult.

---

27 Solo & Snell (1998)

4.1.2 Inefficient duplication – "the game of chicken problem"

Inefficient duplication refers to a duplication in facilities (or more generally in duplication of fixed costs) that is at a net cost to society**.** This can occur as firms race to establish market position.

In such a race, all participants continue to invest until it is clear which firm will win the race. At each stage, investment undertaken at an earlier point in the race can be regarded as sunk, so the competitor will invest up to the point that the required investment is less than the expected value of the prize. The total investment by the winning firm may exceed the value of the market and the resources invested by the losing firms are wasted. In game-theoretic terms this problem can be viewed as a game of chicken.[28]

The problem of a 'game of chicken' has been observed in infrastructure industries at various times and locations. For example:

- in the US in the latter part of the nineteenth century, rail infrastructure was duplicated as companies competed for dominance. Two or more companies would compete, laying parallel routes to serve markets which could not support the resulting capacity. Each company hoped that, whether through completing its line first, or a greater ability to withstand losses, it would remain in the market after the others had withdrawn. Price cutting and destructive competition resulted in losses and scores of railroad bankruptcies in the 1893 Panic;[29]

- it has been argued that duplication of facilities between competing water and gas providers in London in the nineteenth century was similarly wasteful,[30] and could be viewed as resulting from a game of chicken. Another view is that this apparent duplication was often caused by an entrant installing a technically superior system, or otherwise competing successfully, and therefore was not in fact wasteful;[31] and

- this year, as a result of release of new spectrum allowing increased entry into the US wireless phone market, towns

---

28 Kay, John (1993), Foundations of Corporate Success, Oxford University Press pages 46-48.

29 Chernow (1997) p.23

30 Milward (1986)

31 Tynan & Cowen (1998)

which had previously been served by only one or two mobile phone providers are in some cases now served by seven or eight. Many observers believe that the number of providers has exceeded the equilibrium number in many cases. It is likely that several of the providers will be forced to withdraw from the market, having suffered losses and wasted their capital.

The effect on incentives for investment, and thus on service coverage, of a chicken situation are hard to analyse. Sometimes to avoid a game of chicken, no one will enter the market.[32] In other cases, people will rush to invest, and as a result services will be extended more quickly than if free entry is not permitted. If a firm had exclusivity but no legal requirement to invest, the option value of waiting means that even profitable investments may be delayed.[33] (It is perhaps for this reason that most exclusive franchise awards contains service roll-our requirements). Free entry may remove the option to wait and create a rush to invest.

It is in the interests of all firms to avoid wasteful duplication of investment. A number of market solutions to the problem exist. These include:

- market-sharing arrangements and other agreements not to duplicate investments. For example, in Germany, water and gas companies are allowed to enter agreements not to infringe on each other's territories.[34] However in some countries, such agreements would be prohibited by competition (anti-trust) law;

- establishing a reputation for not swerving. This requires repeated games, and reputation may be costly to establish; and

- commitment – in the actual game of chicken, this can be done by throwing the steering wheel out of the window. In infrastructure investment it could be done in a variety of ways, including by sinking sufficient costs early enough, staking something valuable (such reputation) on completing the project, or concluding long term contracts with purchasers.

These market solutions obviously do not always work, or we would not observe games of chicken in business. This raises the possibility that public

---

32 See Kay, John (1993) p.115-116 for an example of this in the UK mobile phone market.

33 Dixit, Avinash K & Pindyck, Robert S. (1994) Investment under uncertainty Princeton University Press, Princeton, New Jersey

34 M Klein, Competition in Network Industries.

policy could improve efficiency by limiting wasteful duplication of infrastructure investment as a result of such games.

However regulators would face a fundamental problem of distinguishing between races that result in duplication which is at a net cost to society, and races that generate benefits through competition and innovation such that there are positive net benefits to society despite some duplication in costs.

Furthermore, games of chicken are not confined to infrastructure, but may be observed in many industries which yield high profits when served by only one or two firms, but which cannot support more than this number. Non-infrastructure examples include the battle between PC (DOS and Windows) and Macintosh formats for personal computers, and VHS and Betamax formats for video recorders. They may also be observed in broadcasting in some countries, and in the civil aircraft market.[35] If the problem is not infrastructure specific, it suggests that public policy makers should consider whether a general solution is desirable, rather than relying on infrastructure specific interventions such as exclusivity.

4.1.3 Inefficient duplication – "the two ice-cream sellers problem"

In a market which can support only two companies, each may each have an incentive to concentrate on the same part of the market, leaving other parts relatively poorly served. This is sometime referred to as the "two ice cream sellers on the beach" problem.

While this problem may not result in unnecessary duplication of fixed costs, it could be considered wasteful duplication in the sense that both firms concentrate in one area of the market while it may be socially optimal to have broader market coverage. This is not therefore a question of whether firms should be allowed to enter the market as a whole, but rather a question of whether firms should be restricted in where they locate in the market. That is, the counterfactual in this case is whether entry should be managed so that entrants are forced to service distinct parts of the market (differentiated by price, quality, or geography).

The theory starts with the observation that the optimal location for single ice-cream seller on a beach is in the centre of the beach, since this minimises the mean distance between the seller and his customers.[36] A second ice-cream seller then enters the beach. She will do best by locating right next to the first ice-cream seller. In this position, half the customers are closest to her, and

---

35 Kay, John (1993) p.115

36 Customers are assumed to be spread evenly along the beach.

half to the first vendor. In any other position, less than half the customers would be closest to her. Turning now to the first ice-cream seller, we see he will not move, since to do so would mean allowing more than half of the customers to be closest to the second vendor.[37]

In this problem, customers' convenience might be maximised by some other positioning of ice-cream sellers. For example, if customers were evenly spaced along the beach, mean journey times to the ice-cream seller would be minimised if one located one quarter of the way along the beach, and the other three quarters of the way along. The market solution under free entry will not produce the socially optimal outcome relative to managed entry, in which the ice cream sellers could be required to service different parts of the beach.

This problem is commonly observed in the petrol station market, where often two petrol stations will locate next to each other (where traffic is greatest) even though motorists might find it more convenient if they were further apart.

In liberalised transport markets, such problems are also quite commonly observed. For example, when the New Zealand government allowed entry into the domestic air services market, the entrant (Ansett) and incumbent (Air New Zealand) scheduled services 10 minutes apart from each other, even though consumers might have preferred more even spacing of flights and hence more choice of flight times.

We are not aware of such problems in piped water and electricity. The most important reason for this is that water and electricity can generally only support one provider in any area. However, even if a market could support two providers of piped water or electricity, we would not expect the two ice-cream sellers problem. This is because the ice-cream sellers problem only likely to exist when:

---

[37] The original paper on these issues was by Hotelling, *Stability in Competition*, Economic Journal 39 (1929). Since then follow up papers have applied different models that vary certain key assumptions, for example Hotelling's assumption of linear transport costs. Overall it is difficult to draw firm conclusions about market equilibrium given different numbers of entrants from the theoretical literature. For example under certain assumptions (quadratic transport costs and Bertrand price competition) the ice cream sellers will locate at the two extremes of the beach to limit the degree of price competition. Clearly in this case duplication does not arise and there would not be any justification for managing entry. We acknowledge the inconclusiveness of the literature but discuss this example because co-location (as illustrated in the two icecream sellers example) may be posed as a reason for managing entry into a market rather than permitting free entry.

- there is a dimension of the product which can take a range of different values (this could be location, or quality where increases in quality require increases in price);

- producers can choose where along this dimension to locate (e.g. centre of beach or end, low quality and low price, or medium quality and medium price); and

- customers are evenly distributed in their preference for different values of the variable in question, and will purchase the product which is closest to their preference.

In piped water and electricity there is generally no choice over location; it must be supplied to the customer's house. Nor will there be much scope to differentiate on quality if, as is often assumed, customers are very concerned that service meet a specified minimum standard, and are generally satisfied once that minimum is met.

That said, it is conceivable that this problem could exist in water and electricity supply. For example, customers might have a range of preferences on some dimension such as reliability of water supply. Those with a sunk investment in storage tanks, for example, might not mind intermittent supply, while others would want continuous supply. It could be that the socially efficient result would be for there to be one cheap unreliable water utility, and another expensive reliable utility. It could also be that free entry would result in two competing utilities of middling reliability and price.

Such problems seem very unlikely in practice. In fact, of the few examples of which we are aware of places which are supplied by more than one water network, the networks are differentiated by the quality of the water they supply. Examples include:

- Hong-Kong also has dual water systems. One supplies potable water, the other non-potable water for toilet flushing;[38]

- a water co-operative in Guatemala City operates two networks, a new one and an old one. The old network supplies customers with water twice a week for a period of two hours each time. The new one supplies water for two hours every day. The tariff differs according to which system a customer is connected to; and[39]

---

38 OECD Pricing of Water Services (1987)

39 Solo & Snell (1998). It is not clear from the source whether the two systems overlap, or serve discrete areas.

- in London at the end of the eighteenth century and start of the nineteenth, many households had more than one company offering them piped water delivery, and some had as many as five. New entrants, such as the New River Company, offered superior service and better quality water than incumbents such as the London Bridge Waterworks, in some cases forcing them out of business.[40]

In these cases, duplication may have occurred, but the choice of product quality suggests that it was not wasteful.

4.1.4 Policy implications

In summary, it is unlikely that inefficient entry and duplication will occur where there is an incumbent and high sunk costs, as in the water and power sectors. It is theoretically possible that where there are strong economies of scale across a segment of the market, free entry could result in inefficient entry, but this is very unlikely in practice.

Inefficient entry and duplication in electricity and water could occur if there was no incumbent, and two or more entrants invested to supply a market which could not support all of them.

The benefits of apparently duplicative entry also need to be considered. In a world where regulation is perfect, duplication in any industry with fixed costs would be inefficient. In the absence of perfect regulation, the benefits of competitive pressure in terms of dynamic efficiency and innovation mean that duplication may not be at a net cost to society. The 'process of creative destruction' which drives dynamic efficiency and technological innovation requires duplication to allow for experimentation, and replacement of outmoded technologies. This suggests that market entry which results in higher industry wide production costs in the short term may not be inefficient over the long term.

Investment that appears to be duplicative may reflect innovation and technological change, or have other benefits. For example:

- the apparent duplication of water and gas networks in London in the nineteenth century allowed companies with better technology to displace incumbents;

- free entry into telecoms, and the significant bypass and duplication which has resulted, is generally considered justified by the benefits of competition and innovation; and

---

40 Tynan & Cowen (1998)

- the benefits of the free entry policy that allowed Ansett to enter the New Zealand air service market in terms of lower fares and improved quality outweigh the possible costs of duplicative scheduling.[41]

Where investments are wastefully duplicative as a result of a game of chicken, the investors will generally bear the costs. Governments do not usually try to protect private investors from the costs of their own mistakes, and it is not clear why this should be a policy objective in the case of infrastructure investors.

This point is worth emphasising. Investors often do lose money when markets become overcrowded, and their investment duplicates that of others. This has happened in infrastructure repeatedly. Canals were bankrupted by railways. US railways competed each other into bankruptcy until they were reorganised by JP Morgan into trusts (cartels and monopolies).[42] These were later hit by the advent of cars and trucks. The American car industry was hit by Japanese imports, fought back, but now the global automobile industry is again suffering from overcapacity and pressure on profits. Similar stories exist in most industries. The invention of the power loom threw weavers out of work, and spawned the Luddite movement. Corner groceries have been outcompeted by supermarkets offering lower prices. Equipment for producing and viewing TV in black and white had to be scrapped when colour came in. This in turn will be superseded with the advent of High Definition Television.

Governments generally allow investors freedom to choose how to invest.[43] Governments recognise that business people have the most information, ability and incentives to make good, profitable business decisions.

If Government tried to protect investors from loss by preventing them from investing in areas where their investment would duplicate existing capacity, the basic processes of the market economy would be subverted, resulting in enormous lose of welfare.[44]

---

41 1996 Commerce Commission Report on the proposed acquisition of shares in Ansett Australia by Air New Zealand.

42 Chernow (1997) p.23-24

43 Generally but not always. The US Government has tried to encourage the move to HDTV by giving broadcasters spectrum for the purpose. Governments in Europe and Japan have at different times protected small retailers by imposing restrictions on the entry of large stores. There are many such examples. But they are the exception, rather than the rule, and in many cases have reduced welfare overall.

44 This can be seen by comparing socialist economies, with Government directed investment into priority areas, and strict rules against duplication of facilities and

The issues in water, sanitation and power investments are little different from those in other areas of the economy. Under free entry, private investors may choose to enter. If an existing monopoly is doing a good job, meeting customers needs at a reasonable price, entry will generally be unprofitable, and no business person will invest. Where business people do invest, it will usually signal that they see an unmet demand from customers, or a way to provide service more cheaply than the incumbent.

In conclusion, this section has shown that there will be a few cases in which free entry leads to inefficient duplication of investment. But this is little different from many other sectors of the economy. The possibility of duplication does not justify any significant restriction on the basic presumption that free entry should be permitted in the water, sanitation and power industries.

An exception to this general rule could be when:

- there is no incumbent in a market; and
- entry to the market is likely to be highly profitable.

It is in these situations that wasteful games of chicken have most commonly been observed in other markets, such as US railroads, the personal computer market, and the video recorder market.

In such a situation government should consider a mechanism such as franchise bidding, which will solve the game of chicken problem, and also address the regulatory issues discussed in Section 5.2. The government should take into account not only the likely costs from a game of chicken, but also the likelihood that market solutions to the chicken problem will emerge. It is notable that in almost all the examples of chicken problems cited, it was not possible for entrants to establish long term contracts with customers. Few people would buy a season ticket on a railroad which had not yet been built, or agree to buy a video recorder which was not yet in the shops.

In the case of networked supply of power, water and sanitation it frequently will be possible to conclude long terms contracts prior to investing. Indeed this is what is observed. For example two companies were in competition to build a gas transmission pipeline between Chile and Argentina. Rather than racing to invest, each tried to collect long term contracts in advance. One secured enough contracts to go ahead; the other did not, and gave up.

Given the costs of Government failure, managed entry, such as through franchise bidding, should only be used in preference to free entry where:

---

unauthorised entry, with market economies, in which entry and duplication are freely permitted.

- there is no incumbent in a market;
- entry to the market is likely to be highly profitable;
- the costs to entrants of assembling long term contracts with potential customers would be high compared to the costs of a Government awarded franchise; and
- the Government has the capacity to run a franchise bid successfully.[45]

In other cases, the presumption in favour of free entry should hold.

4.2 Universal service issues and cherry picking

Governments often want to ensure that infrastructure services are widely available throughout the country. Entry is often prohibited to protect cross-subsidies, which are intended as a way of extending service to areas which might not otherwise receive it, and of providing service at below cost to low income consumers. This continues to be one of the most pervasive arguments for restricting entry.

Examples of exclusivity intended to support universal coverage include:

- in Jamaica, the National Water Commission is legally the sole supplier of water. It is required by central government to supply water to remote and rural areas at below cost. It covers the costs through its charges to urban and industrial customers. This approach is seen as an important part of the country's development strategy;
- in the Punjab in India, farmers are supplied with electricity free of charge. This is intended to promote rural development. Charges to other consumers help to cover the costs of this policy; and
- the UK water industry was charged with a statutory duty to have special regard to the interests of rural consumers. This has been taken to justify continued cross-subsidies from urban to rural water consumers in England and Wales.

[45] This is not to say that franchise bidding should not be used where an existing monopoly is to be privatised. However, such bidding should generally be on the basis of a non-exclusive franchise, so that entry by others is not prohibited.

Requirements on the incumbent to provide universal service by cross subsidising certain customers often cannot be maintained under a free entry policy. If the monopolist is prevented from pricing at incremental cost, others will be able to provide a lower cost supply to the customers who bear the burden of the cross subsidy. This can lead to 'cherry picking' or 'cream skimming', and undermine the profitability of the incumbent monopolist.

Cherry picking and cream skimming are both terms which refer to the behaviour of firms which enter the market to serve those customers that generate the most profits. Cherry picking is likely to occur where:

- the incumbent's tariff is above the cost of supply to different groups, or individuals; and where
- tariffs reflect the marginal cost of supply but do not take into account the risk characteristics of different customers. The expected return from high risk customers is less than from low risk customers. An entrant could cherry pick by supplying only low risk customers.

Cherry-picking is not likely to occur where the incumbent is free to adjust prices in response to entry, even if those prices are initially out of line with cost.

Where investors see a risk that a liberalised entry policy may be reversed in the future, we would expect entry to be less likely in industries with high sunk costs, and long investment lead times. For example, cherry-picking seems more likely by mini-bus operators and call-back companies, than in water and electricity distribution services.

In situations in which the incumbent cross-subsidises, but free entry is allowed, cherry picking is quite common. Examples include:

- in the UK, through water inset appointments. The inset appointments give large users the ability to choose between possible suppliers. Only two inset appointments have been approved by the regulator thus far.[46] However, the threat of inset appointments has driven substantial rebalancing of water tariffs throughout the industry, in order to bring industrial tariff more in line with costs; and
- telecommunications companies in developing countries generally subsidise local charges from international tariffs,

[46] One favour of Anglian Water and against Essex and Sussex Water, and very recently one in favour of Thames Water.

which are well above cost. The huge success of call-back companies is entirely due to this distortion.

Clearly cherry picking will have negative implications for the profitability of the monopolist, and may undermine its ability to meet universal service obligations. For example, in the UK, an interesting tension is starting to develop between the liberalisation of competition in the water industry, and the statutory requirement for the industry to have special regard to the interests of rural consumers.

Cherry picking may also decrease productive efficiency. An entrant can compete even if it is a higher cost producer than the incumbent. For example, assume that it costs an entrant US$0.08 per kWh to supply an industrial estate. The incumbent can exploit economies of scale, so its incremental cost of supply to the same estate is US$0.06 per kWh. However, because of its policy of cross-subsidisation, the incumbent sells to the estate at US$0.10 per kWh. In this situation, the entrant will be successful, increasing the total cost of electricity supply in the country.

While it could be argued that the incumbent could respond by dropping its price in such a situation, in practice this may be difficult. For example, if an entrant threatens to supply only one industrial customer with electricity, the incumbent would need to drop its price to that customer. In many countries, non-discrimination rules would then require the incumbent to drop its prices to all industrial customers. The loss of revenue would be large, and generally could not be recovered by increasing tariffs to residential or other consumers. Given this, the incumbent will often prefer to lose a customer to an entrant than to drop its price. This phenomena can be observed in the international call-back market, where phone companies prefer to lose some customers to call-back companies, rather than reduce international tariffs for all customers.

However cherry picking may also generate benefits. It at least makes cross subsidies more transparent. It generally boosts allocative efficiency by reducing cross-subsidies as the incumbent finds they are unsustainable. Cherry picking can also improve productive efficiency, by creating pressure on the incumbent to improve internal efficiency and reduce costs.

4.2.1 Problems with universal service requirements

There are a number of reasons why a desire to provide universal service will not generally justify restricting entry. These include:

- subsidised universal service may be a poor policy objective;
- there are usually better ways to provide universal service; and
- cross-subsidies often do not work to provide services to target groups.

Each of these points is discussed below. Nevertheless, there are times when restrictions on entry may be justified, either indefinitely or for a transition period. These are discussed at the end of this section.

As noted by Irwin (1998), subsidising infrastructure service tariffs for poor consumers is unlikely to be an effective means of assisting the poorest in developing countries, as many do not receive the network service in question. Even where they do receive the benefit of the subsidy, they might well be better off with a cash subsidy, so that they can choose to spend the subsidy on whatever they need most. For example, some households might choose to spend the subsidy on food or education, rather than water or electricity.

Where governments are committed to providing universal service, exclusivity is not usually necessary to implement such a policy. Competitively neutral methods of achieving the same objective include:

- voucher-like systems for needy households. These are used in Chile for water, and in Romania for district heating; or
- levies on all industry participants paid into a fund to subsidise any firm which supplies services to the target groups.[47]

The disadvantage of these options is that they generally require higher institutional capacity than a system of exclusivity supported by cross-subsidies.

On the other hand, it must be borne in mind that attempts to provide universal service by supplying target groups below cost frequently fail. This is a problem for both public and private operators. For example:

- in Trinidad, the water utility was publicly owned and operated until 1996. Only 11% of the population had a continuous water supply. This 11% was predominantly urban and well-to-do. Customers in South Trinidad, a largely poor and rural area, seldom received water more than three times a week;
- in many African cities large proportions of the population do not receive piped supply, and rely on private provision by water vendors. These vendors charge much more per cubic metre supplied than the formal utilities. For example in Mandera, Kenya, 90% of the population receive water from

[47] Irwin, *Price Structures, Cross–subsidies and Competition in Infrastructure* Public Policy for the Private Sector 1997.

private vendors, and spend on average more than 30% of their income on water;[48] and

- in Buenos Aires, it is reported that the concessionaire is not doing as well as was first hoped in expanding service to unserved areas of the city. The concessionaire is reluctant to expand service, since the tariff structure means it will generally lose money by doing so.

In many developing countries a monopoly supplies water or electricity at below cost to higher income customers, while slums get no service. This approach is often institutionalised, with utility master plans simply not including any provision for slums. In many cases poor consumers rely on high cost water carriers, and do without electricity, or connect illegally. In this situation, free entry could hardly make coverage and equity worse, and could help.

Examples of entry helping to expand service to target groups include:

- the Orangi Pilot Project in Karachi. Orangi is a slum area in Karachi, which had extremely poor sanitation. Community groups developed and implemented successful low cost piped sanitation systems. If the incumbent utility had had and enforced exclusive rights, this could not have happened;[49]

- similarly, the Indian city of Ahmedebad is developing successful 'slum networking' projects. These bring low cost network services into slum areas, supported by local businesses and community groups. Again, they would not be possible unless entry was permitted; and

- in Maputo, Mozambique, wealthy residents with (subsidised) piped water supply run informal plastic-pipe systems to supply water from their connection into low income areas. They sell the water for much more than it costs them, and at prices which can cause real hardship to their low income customers. These informal systems are illegal, and the price-gouging of the poor by the well-off, using subsidised water, is generally regarded as inequitable. On the other hand, at least this system provides the poor areas with a piped supply. Enforcing the prohibition on such systems would remove a valuable service. Conversely, removing the prohibition would

---

48 Lewis, Maureen and Miller, Ted R. (1997) *Public -Private Partnership in Water Supply and Sanitation in Sub-Saharan Africa* Health Policy and Planning, (2) 1.

49 Solo and Snell (1997)

help to bring the systems out from the underworld. Since the pipe-systems are cheap to install, it could even open the way for competition between suppliers to the slums.

A similar example can be found in the electricity sector in the Yemen where the public utility meets urban electricity demand, but does not supply rural towns and villages. There is little prospect of significantly improving the proportion of rural grid-connected households in the next two decades, but small scale private sector operators have demonstrated a willingness and innovative approaches to meeting the demand. These small-scale providers range from individual households generating for their own use and supplying neighbours, to larger operators supplying up to 200 households and operating sophisticated billing systems. The result is that rural household electricity use in the Yemen is remarkably high compared to other middle income countries. [50]

4.2.2 Policy implications

The key policy implications of this discussion are that:

- subsidised universal service is often a poor policy objective. Where governments do not have such an objective, there is no need to restrict entry on this account;
- in many places, attempts to provide universal service through cross-subsidies supported by exclusivity have failed to achieve their objectives. In such circumstances, restrictions on entry should usually be removed; and
- where universal service supported by exclusivity is working, it will generally be better to switch to a competitively neutral mechanisms to support universal service. Once this has been done, restrictions on entry will no longer be justified.

In some cases, Governments will be committed to universal service, but will not have the capacity to administer competitively neutral mechanisms. In such cases, restrictions on entry may be justified, if there are good prospects that this approach will achieve the objective.

Where Government decides to move away from a policy of cross-subsidisation, consideration could be given to continuing to limit entry for a transition period. Otherwise, higher cost suppliers may be able to enter the market and take advantage of price distortions before the incumbent is able to rebalance tariffs. If the entrant has sunk costs, and can lock in demand

[50] ESMAP (1991).

through long term contracts, then this could result in a higher overall cost structure for the industry.

4.3 Exclusivity and involving the private sector

Exclusivity is often granted to encourage private sector participation in an infrastructure sector. Exclusivity clauses have been a very common component of moves to involve the private sector (e.g. the granting of concessions, leases, or assets sales) in telecommunications, water and electricity. For example:

- in Jamaica, when the telecommunication utility was privatised in the 1980's a 25 year exclusivity clause was included;
- in South Africa in 1997 an exclusivity period of five years was extended to Telkom for local long distance and international services. Under the agreement Telkom was obliged to meet a series of service extension and quality improvement targets, and is subject to financial penalties if it fails to meet those targets;[51] and
- in Buenos Aires, as in most places which have involved the private sector in water supply, the water concession gives the concessionaire the exclusive right to supply water within the concession zone. In other words, competitive entry is prohibited.

An example of the value of exclusivity clauses is given by Hong Kong. In 1998 the Hong Kong government paid US$866 million (HK$6.7 billion) to Hong Kong Telecom to buy out the last eight years of its exclusive international license.[52]

It is important to unravel why exclusivity is granted. In many cases it is to preserve cross-subsidies. This was the case in South Africa telecommunications for example, and it may have been a factor in Buenos Aires water.

If the private investor is not required to preserve cross subsidies, the main benefit of exclusivity is to remove the demand risk created by competitive entry. As discussed in Section 4.1, there may also be occasions in which it can increase efficiency by reducing the risk of inefficient duplication of facilities.

---

[51] Financial Times Business Report: Technology (Q3:35) 13 February 1997.

[52] 31 March 1998 Reuter News Service – Far East Reuter Textline (Q2:60). March 1998 Exchange rate US$1:HK$:7.74 as at March 1998.

4.3.1 Circumstances in which exclusivity may be important

If a government is intent on maintaining cross subsidies (and will not adopt a competitively neutral method for cross subsidisation as outlined in Section 4.2 above), then it is very likely that an exclusivity clause will be required.

If exclusivity is granted to remove demand risk, the value of the exclusivity clause will vary across industries. We distinguish between these three cases:

- competitive entry is unlikely, such as in water distribution (with no cross subsidy requirements);
- competitive entry is likely, for example in bus services or cellular phones, and consequently exclusivity is valuable to the incumbent; and
- competitive entry will only result from a positive decision by Government. For example planning approvals and land acquisitions for urban rail, toll roads or airports, are often so complex that new entrants require special legal support from Government.

In the first case, where exclusivity has a very low value, not granting exclusivity should not affect the willingness of the private sector to invest.

In the second case where exclusivity is valuable, it is also very also costly to the economy as a whole. Governments need to weigh up whether the benefits of private participation outweigh the costs of the hidden subsidy, and also whether exclusivity is the least cost way to deliver the subsidy. Often there will be better options, such as accepting a lower sale price for the asset.

In the third case, the Government has private information about its likely future action which is highly valuable to an investor. A toll road to the airport may be worth building, but not if within a few years a new rail line will opened to compete with it. [53]

If the Government does not envisage facilitating a competitor, it needs a credible way to signal this to the first investor. If it does not signal its intention credibly, the investor may assign too high a probability to the Government promoting a competitor in the future, and not invest, (or demand higher fares, or make a lower payment for the franchise right). A good way for the Government to credibly signal its intention not to promote a competitor is to grant an exclusive franchise.

---

[53] This could be explained as a 'market for lemons' problem. See Akerlof, George, "*The Market for 'Lemons': Quality Uncertainty and the Market Mechanism* Quarterly Journal of Economics (August 1970).

It is important to bear in mind however, that the private sector often does invest in infrastructure without exclusivity. For example:

- water supply to London in the eighteenth and nineteenth centuries was developed largely by private firms operating under non-exclusive franchises;[54]

- investors in US railroads at in second half of the nineteenth century seldom had exclusivity rights. Sometimes, other companies would start to lay competing lines solely to extort payment from the first company for abandoning the line.[55] Nevertheless, huge sums were invested: $300 million in the 1840s, and $840 more in the 1850s, mostly from foreign investors. These sums translate to billions of dollars in today's money. By the 1890s, roughly 60% of the companies on the New York Stock Exchange were railroads;[56] and

- in the present day, as Box 4-5 illustrates, private companies are risking billions of dollars in satellite communications infrastructure which will serve developing countries without exclusivity.

It is worth noting that it is often the lack of Government restrictions on entry and other aspects of business which encourages private investment. This year the US company Formus Communications invested substantial sums to purchase radio spectrum in New Zealand. Formus plans to provide a wireless Local Multipoint Distribution System (LMDS). This is a new broadband technology capable of providing phone service, subscriber-TV and high-speed internet access. The key reason for choosing New Zealand for one of the company's first investments in this area was that country's extremely liberal policies on entry in the telecommunications market**.**

54 Tynan and Cowen (1998)

55 Chernow (1997) p.20

56 Wasserstein (1998) p.33-34

Box 4-5 New satellites – huge infrastructure investments with no exclusivity.

Iridium is launching a network of 66 Low Earth Orbiting Satellites (LEOS), at an estimated cost of US$25 billion. From September this year the satellites will provide voice, fax and limited data transmission capabilities without the need for any ground infrastructure. Consumer equipment, such as mobile phones, will communicate directly with the satellites. This will revolutionise communication across much of the developing world. This massive investment in communications infrastructure is being made without any exclusivity. While Federal Communications (FCC) and International Telecommunications Union (ITU) rules protect the spectrum Iridium will use, the market itself is wide open.

Freedom of entry is demonstrated by the number of other companies in the market. Globalstar is launching a competing LEO system at a cost of US$2.7 billion. Teledesic is investing over US$9 billion on a broadband satellite system which will compete with Iridium in some markets, as well as providing video conferencing and other high bandwidth capabilities. Skyborg plans to spend US$3.5 billion on another broad band system.

Sources: www.sat-net.com www.skyborg.com www.Teledesic.com Trinidad Sunday Guardian 7 June 1998 page 3, Barron's online June 15 1998

4.3.2 Policy implications

As noted above it is likely that exclusivity clauses will be required if governments wish to retain cross subsidies, and are not willing to develop a competitively neutral way to implement them.

In the absence of requirements on the operator to cross subsidise some consumers:

- where the likelihood of entry is low, governments should explore whether private investors really value legal exclusivity, and aim to avoid restricting entry. Often not granting a legal monopoly will have public relations benefits which add value for the investor by making the transaction more politically sustainable;

- if entry is likely, the benefits and costs of exclusivity must be carefully evaluated, in order to assess the lowest cost ways to meet the government's objectives for the sector. Often there will be better options than granting exclusivity. For example

governments may accept a lower sale price or grant limited guarantees to cover certain risks; and

- when intervention by Government would be required to allow further entry, and Government does not intend to allow such entry, it will often be best to grant exclusivity.

Exclusivity is common in private infrastructure deals, so investors may be unduly nervous about deals which do not involve exclusivity. Another way of saying this is that, given a lack of experience with non-exclusive transactions, investors may mistakenly overprice the risk of non-exclusivity. Given information asymmetries between investors and the governments, investors may also interpret the refusal to grant exclusivity as a signal that competition is more likely than it really is.

Where governments cannot persuade investors to price non-exclusivity more appropriately, a good compromise may be to offer exclusivity for a limited period. For example, short five or seven year exclusivity periods have become common in telecommunications, for example in the South Africa example mentioned above. Offering such terms may give investors sufficient comfort, without doing undue harm. Again, it will be a case-by-case decision.

4.4 Scheduling and routing

In some infrastructure industries the timing of delivery to certain parts of the network is of key importance. This is particularly significant in industries such as transportation or telecommunications where storage costs (or costs of delay) are high, and services are not homogeneous but have a value that is determined largely by the route. Where this is the case, routing and time-tabling become additional constraints that need to be taken into account in the network optimisation problem.

Scheduling and routing problems could arise as a result of:

- ongoing revision of schedules in an attempt to reach a point in the network ahead of competitors – the problem of the three ice-cream sellers on the beach; and

- difficulty in co-ordinating timing and establishing optimal delivery schedules across a complex network.

4.4.1 Schedule jockeying – unstable market equilibrium

In industries where timing is the key determinant of the consumption decision, free entry may not arrive at a stable market equilibrium and may not generate beneficial competition.

This can be illustrated by reference to the game theoretic example outlined in Section 4.1.3 above of two ice cream sellers on a beach co-locating.[57] When there are three ice cream sellers rather than two, the sellers will not settle at a stable position on the beach. The sellers will continue to re-locate, each one adjusting according to the position established by the other two in an attempt to maintain or improve position in the market relative to the competitors. The deciding factor in consumers' demand decision is proximity, so location is key to market share, and thus to revenue and profits.

This outcome has occurred in some cases where there has been free entry in urban bus industries. Scheduling is a problem because waiting time is so dominant in passengers' travel decisions that customers will almost always take the first bus that comes. This leads to ongoing revisions of schedules as each bus company seeks to schedule its services just ahead of the competitors' services. Furthermore, it means there is little incentive to compete strongly on price or quality.

For example, problems emerged following the deregulation of urban buses in Britain during the 1980s. All urban bus services (except in London) were deregulated and privatised. Where competition occurred entrants competed by scheduling services along the same routes for very similar fares. Destructive competition emerged as routes were swamped by incumbents responding by trying to schedule services so frequently that entrants could not get enough passengers to be viable.[58]

Scheduling problems such as this are most likely to occur where the demand decision is dominated by timing considerations and there is:

- a homogeneous product with little value attached to brand;
- low sunk costs; and
- low costs of changing timetables.

A market solution could exist to these problems. If competition over routes and schedules is unsustainable, one could assume that the competing firms would have an incentive to define geographical monopolies. However, the question remains whether:

- transaction costs and competition law would allow such a market solution to occur;

---

57 We note that as (discussed in footnote 37), the literature is inconclusive about the impact of the number of players on the stability and location of market equilibrium.

58 Klein, Daniel and Moore, Adrian, *Schedule jockeying and route swamping: bus markets in Britain need kerb rights* Institute of Economic Affairs (June 1997).

- the solution would be sustainable since new companies which were not part of the original agreement could always enter an established company's area; and

- the process of arriving at the market solution, and indeed the equilibrium outcome, would be less costly to society than restrictions on entry and competitive award of monopoly franchises.

In industries such as urban transport, the typical regulatory response has been franchising of monopoly rights to groups of routes. Where there are low sunk costs, frequent tendering for monopoly franchises may offer a good solution. While it prevents free entry for the period of the franchise, it offers regular opportunities to firms to enter the market, thereby maintaining competitive pressures. This has been followed in the Hong Kong bus industry, for example.

We are not aware of any instance when schedule jockeying has been a problem in network water supply or electricity distribution. It is difficult to conceive of a situation in which it could occur.

4.4.2 Optimal delivery schedules in a complex network

A policy of free entry, and the existence of a number of competitors in the market, can increase the complexity of optimising delivery schedules across a network. As the number of firms competing for use of the network increases so does the difficulty of efficiently adjusting delivery in response to changing demand and supply conditions.

Optimising delivery schedules becomes important when:

- the good or service is not storable, or the costs of storage are very high; and/or

- the units are not homogeneous such that value of any unit or service is a function of delivering it to a particular customer or specific points in the network at certain times, for example in transport services, or communications.

In industries such as aviation, railways and telecommunications, both of the above two conditions apply. The costs of delay are high, and the service is differentiated by its supplier, its route and schedule. For example, a service by one airline may be valued differently from the same service supplied by another airline (because of quality of service, safety concerns etc.). Free entry, and the presence of competitors will clearly increase the complexity of optimising delivery schedules in such a system.

A market solution could work through defining rights over points in the network, allocating those rights, and then allowing trading to occur. It should ensure that rights are allocated to the highest value user. This would be a complex market and, as Klein argues, decentralised bargaining may not generate an optimal timetable or allocation of routes through the network.[59] A computer based auction system that simultaneously generates the optimal prices and set of paths through the network may be required to achieve optimality. Such a system could be highly complex, and the costs to market participants of understanding and using the system could also be high.

It has long been argued that a market system to allocate landing slots or track rights could be applied in airports and railways, but to date one has not been implemented. It has been considered in the UK for airport landing slots and railways, and in Sweden for railways.[60] Markets do exist for telecommunication capacity, but as yet these use long term contracts, and do not have any spot-clearing mechanisms, making the resulting allocation of capacity less than perfect.

The typical regulatory response in railway and airports has been to use variations on administrative processes for allocating rights. These administrative processes are costly and unlikely to generate an efficient outcome. Although policy makers have not embraced market solutions to allocation problems in these industries, even imperfect market solutions may be better than existing administrative approaches. Courier companies, trucking businesses and deregulated telecommunications companies provide examples of imperfect but highly workable solutions to many of these problems.

Nevertheless, it is sometimes argued that scheduling and routing problems are a market failure, which justify restrictions on entry to ensure that administrative co-ordination can be achieved. Difficulty in obtaining slots at Heathrow airport, for example, is a major barrier to entry to European civil aviation markets. How do these argument apply to water, sanitation and power?

Water is storable at low cost, and is generally a homogenous product. Wastewater is less homogenous, in that different industrial process can produce different wastes, which may need to be treated in different ways. However it also easily stored. Scheduling and routing issues are unlikely to be a major concern in water and sanitation, and would not justify restrictions on free entry.

---

59 Klein, Michael, *Competition in Network Industries* Private Sector Development Department, World Bank.

60 Klein, Michael

Electricity itself non-storable. Electrical energy can be stored by converting it into other forms of energy. Rechargeable batteries are an example, but this method of storing energy is so expensive that is viable only in specialised uses. Power utilities in several countries use 'pumped storage'. Off-peak energy is used to pump water uphill into a reservoir. At peak times, the water is allowed to flow down again, passing through turbines and generating electricity. However, the cost of storage is very high. In addition to the cost of the storage facilities themselves, much energy is used in the process. It is only the large price differential between peak and off-peak power which makes this approach viable. This is in marked contrast to water, where the product itself does not diminish in value by being stored.

The location on an electricity grid of suppliers and demanders of energy also matters. As power flows through a network, transmission capacity surplus or deficit is a function of total flows throughout the system. In an open access system it is difficult therefore to define capacity or access rights to a power system. Long term un-tradable contracts are not an efficient way to match supply and demand. This problem can be resolved through a central dispatch system which optimises system flows, and instantaneously matches supply and demand such that the highest value user is supplied.[61]

Various forms of central dispatch systems for electricity have been set up in a number of countries, with considerable success. In Argentina, a competitive market was introduced in 1992, in which bids are based on audited costs of power plants rather than generators' price bids. Klein notes that the switch to a private competitive system quickly resolved all the urgent problems of power shortages and created a situation of temporary excess capacity as new generating firms rehabilitated and operated existing plants.[62]

The value of central despatch optimisation may justify some government intervention in the industry. But it should be noted that market participants have incentives to maximise gains from trade by creating efficient trading systems. Box 4-2 provides an example of how markets can reach efficient despatch solutions, even when these are not designed into the market structure from the start.

---

61 Klein, Michael

62 Klein, Michael

Box 4-2 Norwegian over-and-under electricity pool

The Norwegian electricity industry contains around ten competing generators, and several hundred electricity retailers. There are long term contracts between the generators and the retailers. Exclusive and non-tradable long term contracts are not the most efficient way for generators to interact with retailers. This is because there will be times when one generator is contracted to supply a retailer, but another generator could in fact do it more cheaply.

Recognising the gains from trading in this situation, the generators initially created a market in which they could trade between themselves. This means that if Generator A is contracted to supply electricity to a customer, but Generator B could produce the electricity more cheaply, Generator A can meet its contractual obligations by purchasing the electricity from B in the spot market, and then on-selling it to the customer.

The market has subsequently evolved to include demand side bidding. This means for example, that if a customer has a long term contract, and does not want to use the full amount for which it has contracted, it can sell the portion it does not wish to use in the market, effectively adding it to supply. Customers can also purchase electricity in the spot market.

It has recently been extended to incorporate trading with Sweden and Denmark. The result is that the Norwegian system has evolved into a market with efficient despatch of generation (99% hydro) that allows market participants to optimise storage level in their reservoirs.

This suggests that while free entry might initially produce a system with inefficient despatch, as the potential gains from efficient despatch grow, it is realistic to expect market participants to develop an efficient despatch system, to allow them to realise those gains.

4.4.3 Policy implications.

Free entry is unlikely to create scheduling and routing problems in the water and sanitation industry because:

- the high sunk costs and difficulty of relocating means that ongoing repositioning is unlikely to be a problems;
- water is storable so the problem of matching timing of delivery and capacity is less significant; and

- water is a homogenous product (wastewater may be less so) and the value of a cubic meter to the consumer should not be greatly affected by its supply route.

Scheduling is however an issue in the electricity industry because in electricity timing matters, and the capacity in the system at any moment is a function of all flows throughout the system. This is only a problem on complex inter-connected systems. While free entry by itself is not likely initially to result in a complex inter-connected system, in time it could result in such a system. However, for that to occur, either the Government or the market would need to resolve the common carriage problem. The same mechanism used to resolve the common carriage problem could also be used to resolve the despatch problem. That is, either the government or the market could determine the despatch mechanism concurrently with or subsequent to determining the common carriage arrangements. Therefore, routing and scheduling problems do not justify restrictions on free entry.

In summary, scheduling and routing problems will not generally be a reason to limit free entry for water and electricity. Some policy interventions may be justified for industries with low sunk costs, such as solid waste collection and buses. Bidding exclusive franchises is one option. Other options should also be considered, such as simply requiring operators to publish and adhere to timetables.

4.5 Environmental quality and safety outcomes

To the extent that infrastructure providers have an incentive to reduce cost or increase revenues, they will have an incentive to give inadequate weight to environmental values. For example, many water utilities over-abstract acquifers, because this can be the cheapest way to supply water to people. If personal injury law does not provide an effective deterrent, utilities may also skimp on safety. For these reasons, Governments often impose environmental and safety obligations on infrastructure providers.

In some instances environmental and safety problems may become acute if free entry is allowed and competition develops. This could result from two factors:

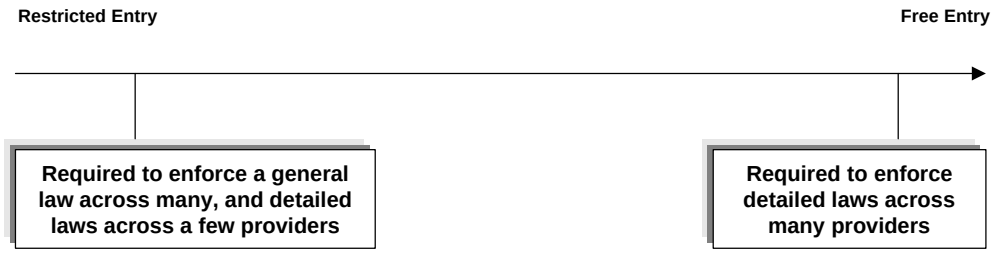
- as the number of providers increases, the job of enforcing environmental and safety standards becomes harder; and

- competition increases the pressures on providers to cut costs and win customers. This increases their incentive to cut corners on environmental quality and safety.

4.5.1 The enforcement problem

In many developing countries enforcement capability is limited. As illustrated in Figure 4-1, free entry changes the nature of the enforcement problem.

Figure 4-1 : Enforcement requirements under restricted entry and free entry



Under restricted entry, enforcement agencies only have to police a broad requirement across the whole industry: that unlicensed providers do not enter the industry. Policing of detailed safety, quality and environmental regulations is limited to the few licensed providers. Where a free entry policy results in an increase in the number of providers, the enforcement agencies must police specific laws across a greater number of firms. In some instances this may be beyond the capability of those agencies.

4.5.2 Effect of increased competitive pressure

Where specification, monitoring and enforcement of environmental and safety standards are poor, the issue of whether to allow entry can be analysed in terms of high and low powered incentives.

The theory is as follows[63]: Imagine a worker whose output has some attributes which are easily monitored, and others which are hard to monitor. The worker can be given either high or low powered incentives. High powered incentives are ones which strongly motivate performance. Piece work rates and success fees are both examples of high powered incentives. Low powered incentives are less performance related. For example, a fixed monthly salary is a low powered incentive compared to piece rates or bonuses based on profitability.

63 See Milgrom and Roberts, Economics Organisation and Management 1992 Chapters 7 and 12 for further discussion.

At first sight it might seem more efficient to always use higher powered incentives. In fact, this can be counter-productive. If some of the desired attributes are hard to measure, high-powered incentives cannot be provided for these attributes. Providing higher powered incentives on the other attributes will skew the worker's efforts toward the measurable outputs, with a loss on the unmeasurable outputs. Lower powered incentives will not motivate such hard work on the measurable outputs, but will also not cause effort to be diverted from desirable but hard to measure outputs. Often the overall result from the lower powered incentives will be better.

To give an example, the number of studies reviewed by a consultant in preparing a report is relatively easy to measure, while the quality of the consultant's analysis of those studies is harder to assess. The client is interested in maximising both the number of studies reviewed and the quality of the analysis. A high powered incentive scheme, such as a low fixed fee coupled with a bonus for each study reviewed, would probably result in poor analysis. The client can do better by simply setting a fixed fee for the whole project, even though this does not motivate the consultant to review so many studies.

Turning to utilities, we see that they produce a range of desirable outputs and characteristics, which vary in how well they can be measured. For example, profit is easy to measure. Where capacity to monitor and enforce environmental and safety standards is low,[64] environmental and safety characteristics are obviously hard to measure.

In a monopoly environment, the pressure to make profits is often lower than in a competitive environment. When competition is introduced, the pressure to increase revenues and cut costs gets stronger. In other words, the power of the commercial incentives increases. However, the power of the environmental and safety and incentives remains as low as before. In this situation, the industry's behaviour can be skewed away from environmental and safety objectives. It is possible that the loss on these objectives would outweigh the gains from increased commercial efficiency.

Box 4-3 illustrates a problem of enforcement caused by free entry, and arguably also a skewing of incentives toward commercial performance at the cost of safety.[65]

---

64 Enforcement capacity could often be increased by spending more on the enforcement agency, but in many countries Government budget constraints prevent this.

65 As well as a time-tabling problem (see Section 4.4)

Box 4-3: Law enforcement under free entry in the Jamaican bus market

In Jamaica, largely free entry was allowed into what had been a Government urban bus monopoly in Kingston. This resulted in numerous under-capitalised owner-operators entering the industry. The bus operators would race along city streets to beat rivals to the next stop. They even took shortcuts, going off road and through public parks. The buses themselves were often under-maintained and overloaded.

In this case, strict enforcement of the traffic code would have stopped the problem. But the number of offenders, and their strong incentives to offend, overwhelmed Jamaica's enforcement capacity. The problem was probably worsened by the fact that the businesses were small, undercapitalised and only marginally profitable. Since they had little to lose, it was hard to apply effective sanctions.

The Government's approach to the safety and quality problems caused by numerous small operators has been to encourage the industry to merge into a small number of larger companies, each of which was then awarded an exclusive franchise to an area of city. Whatever the merits of this approach overall, casual observation suggests that it has lead to greater discipline and fewer safety infringements in the industry.

A similar example is provided by cable TV industries in many developing countries. Often cable operators are small scale entrepreneurs who pirate programming from a satellite feed, and distribute it through a small neighbourhood network, in breach of the intellectual property rights of the programme owners. Because their operations are small scale, with fairly low sunk costs, they are very hard to police. Again, the solution in Jamaica and elsewhere has been to award a limited number of exclusive geographic franchises. Each franchise holder then has an incentive to stop pirates operating in their area, while the Government is able to ensure that the larger operators comply with the law.

4.5.3 Industries where enforcement problems are most likely to occur

In general, the problem of enforcing environmental, safety and other legal obligations will be greater the lower the sunk costs of the industry, the smaller the efficient scale of operations, and the poorer the government's enforcement capacity. Table 4-1 provides a judgmental categorisation of various infrastructure and related industries.

Table 4-1: Enforcement problems

Likely	Bus transport
	Water trucking
	Night soil collection / cesspool emptying
	Bottled gas supply
	Cable TV
Less likely	Small scale water abstraction
	Packet sewage treatment
Least likely	Electricity generation and distribution
	Telephone network
	Water distribution
	Gas distribution
	Railways

Even industries in the 'least likely' category will infringe environmental and safety obligations if there is no enforcement at all. In all cases, the best solution will be to effectively enforce safety, environmental and other obligations on all participants in the industry. However, in cases where enforcement capacity exists but is limited, limiting entry can assist enforcement. In some cases, the benefits of better enforcement will outweigh the costs of limiting entry.

4.5.4 Policy implications

Where enforcement capacity is low, the costs of greater infringement of safety, quality, and environmental regulations may in some cases outweigh the benefits of increased competitive pressure. Restrictions on entry may be justified, especially in those industries such as public transport where the risks of enforcement problems are highest.

In electricity generation our judgement would be that environmental concerns will only occasionally justify restrictions on entry. Generators are usually fairly few in number, and easy to spot. An environmental agency which can police the generation stations of the incumbent will usually also be able to police the generators of a few entrants.

Safety concerns exist in all parts of an electricity network. However, those who are hurt by or at risk of unsafe electricity installations usually have a

direct incentive to enforce safety provisions themselves. So long as personal injury law functions reasonably well, safety concerns are unlikely to justify limitations on entry. If there is a risk of under-capitalised operators entering the industry, a requirement to have third party liability insurance could be justified.[66]

In the water industry, safety is much less of a concern, but environmental issues can be very important. Each situation needs to be assessed on its own merits. There may well be cases in which:

- an acquifer or important marine habitat would be seriously damaged by uncontrolled water abstraction or sewage discharge; and

- such abstraction is much more likely if anyone who wants to is allowed to sink their own well or operate a sanitation business.

In such cases restrictions on entry may well be justified. This is particularly the case for water abstraction. Creating a monopoly abstractor partially internalises the environmental externality. That is, because the water company is the only one using the acquifer, it suffers more of the costs of over pumping, and therefore will have both the incentive and the ability to limit over-abstraction. Note however that the internalisation is only partial. Problems of subsidence, for example, will continue to affect people other than the abstractor.

Restrictions on entry will not be justified if any of these conditions hold:

- the environments are not under pressure;

- enforcement capacity is good; or

- enforcement capacity is so low that the incumbent is already seriously infringing environmental standards, and there is little prospect of this changing.

[66] This would be similar for the requirement which exists in many countries for motor vehicle drivers to have third party liability insurance.

5 Associated systems

A policy of free entry would have political implications and effects on other areas of regulation. This section explores the likely impact of free entry on other technical economic and political systems. We outline the relationship between free entry and:

- technical standards regulation and the development of integrated systems;
- broader economic regulation; and
- the politics of monopoly and free entry.

5.1 Technical standards and integrated systems

If a free entry policy is to be adopted it becomes important to ensure that technical standards regulation:

- promotes the development of systems that can be connected to provide an integrated system in the future;
- does not create unnecessary barriers to entry; and
- and does not limit the ability of firms to innovate.

The question of what technical standard to adopt is a clear illustration of the 'battle of the sexes' game[67]. This game is characterised by the following preferences and interaction. Two people are planning to meet for an evening. John prefers to go to a knitting evening, and Sue to go to a sumo wrestling evening, but each would rather be with each other than apart. The worst possible outcome is that John ends up at the wrestling demonstration and Sue at the knitting. Each knows the other's preference, and must plan their actions according to their expectation of how the other will act. Co-ordination could assist the couple in arriving at the most preferred option. It is not clear, however that co-ordination will always succeed especially given the difficulty of assessing the relative strength of each person's preferences, and gains from co-ordination. It would be possible to always ensure the preferred outcome is reached (i.e. going to the same event) if a decision can be made and enforced (either by John or Sue or by a third party).

In the case of technological standards, a battle of the sexes problem could occur where the incumbent, company A, has a 50 Hz electricity system that

67 Kay, John (1993)

supplies the main urban centres in the north of the country, and is expanding rural supply. Company B is proposing to set up to supply the south using a 60Hz cycle. Each company wants to use its preferred technology and system, but their most preferred outcome is that compatible systems are used and an integrated network can be developed in the future. The worst outcome is that both firms use technology and develop systems other than that which they prefer.

Examples of where technical standardisation has failed to occur include:

- Jamaica has a 50hz interconnected electricity grid which now covers most of the country. Before the grid was so widespread foreign owned bauxite companies installed their own generating systems based on a 60hz cycle. The peak times on the bauxite systems differ from overall system peaks. It would be highly beneficial in terms of cost and security of supply to connect the bauxite generating plants to the grid. However, the incompatible technical standards prevent this;

- in several parts of the world, different organisations built railways of different gauges. The networks extended to the point that they connect, and it would be highly beneficial for passengers and goods to be able to move from one system to another without changing trains. However, the different gauges make this impossible. Places where this problem exists include Australia, Spain, and the railway connection between Thailand and Malaysia; and

- in telecommunications, the failure to establish a common standard internationally for phones jacks means that at times, wiring and plug design are often not compatible[68]. This can prevent users connecting to different systems. The same problems applies with incompatibility of power points.

5.1.1 Conditions when technical standardisation is most important

Ensuring technical standardisation is likely to be most important where:

- the costs of converting from one standard to another are high – plug adapters are cheap, so lack of international standardisation in electricity outlets is annoying, but does not have high costs. Electricity can be transformed from one voltage level to another quite cheaply, so lack of voltage standardisation is also not overly costly. However,

68 Negroponte 'RJ- 11'. Wired Magazine. Issue 6.04 April 1998.

transforming electricity from one frequency to another is more costly, and therefore the gains from standardisation are higher. In water however, the costs of connecting systems with different specifications are typically relatively low so the gains from standardisation are much less;

- storage costs are high as a percentage of value – where delays do not matter much, for example in transporting commodities, the cost of converting from one system to another will be lower for example, than in telecommunications, where 'storing' a message until it could be converted to a different format would destroy the ability for two people to have a conversation – the essence of a telephone service. In the electricity industry, the difficult in storing electricity makes the value of electricity vary according to the load on the system, and thus increases the value of connecting systems with different load profiles. Water by contrast is easily storable, so the benefits from interconnection, while still positive, are lower;

- network externalities are important – in some infrastructure industries, the value of the network increases with the number of people using the network. For example, the first telephone was worth nothing, since there was no one to call. Once the second telephone was made, both instruments had value, and their value increased as more and more people were added to the network, and the number of people who could be called increased. It has been argued that value in fact increases exponentially with the number of people connected to the system. There are therefore great gains in connecting telecommunications networks. In contrast, the value to one consumer of piped water varies very little with the number of other people who also have piped water. Network externalities, and therefore the value of interconnection, are lower in water;

- sunk costs are high as a percentage of total costs – if costs are not sunk, it will be cheaper to change one part of a system to make it compatible with other parts. An airline wishing to standardise on Airbuses to reduce its inventory costs can easily sell its Boeings at full value. In contrast, a railway would have to write-off a considerable portion of its existing investment in changing from one gauge to another; and

- technological change is slow – this is similar to the sunk costs issue. Where technological change is rapid, the need to upgrade will often create an opportunity to standardise across a previously diverse system. It may also imply that new ways

to translate easily between standards are more likely be developed.

Table 5-1 below lists the industries in which these factors are likely to be important.

Table 5-1: Technical standardisation

Factors that increase the importance of technical standardisation	High storage costs	High costs of converting	High network externalities	High sunk costs	Slow change in technology
Relevance to industries (in decreasing order to importance)	Telecoms Electricity Transport Gas Waste-water Water	Electricity Telecoms Gas Transport Waste-water Water	Telecoms Transport Electricity Gas Waste-water Water	Waste-water Water Gas Transport (rail) Electricity Telecoms Transport (road)	Water Waste-water Gas Electricity Transport Telecoms

5.1.2 Market solutions

Market solutions are possible, and can in fact be seen as a driving force for innovation.

As suggested by the analogy of the battle of the sexes the outcome can be improved if co-ordination between industry participants is possible. In the case of technology standardisation it is in the interests of all industry participants to commit to a certain standard. This could be done, for example through a voluntary association responsible for setting pipe diameters, and junction protocols in the water industry.

Market solutions can however cause problems. For example:

- a technically superior standard may loose to an inferior standard. Some commentators have alleged that this was the case in the battle between IBM compatible PCs and the Apple

Macintosh, or in the battle between the VHS and Betamax standards for video recorders.[69]

- there may be considerable delays in establishing a standard. This could result in a costly game of chicken as described in Section 4.1.2, and leave some consumers with incompatible equipment. The time taken to establish a standard for 56kbps modems may be an example of this; and

- different standards may become entrenched in different areas, for example US and UK television formats.

The potential costs associated with market solutions has led to calls for Government intervention to establish open, industry wide standards. However, this can also be extremely costly.[70] Unless carefully designed, technology standards can grant a competitive advantage to certain companies within the sector, become a barrier to entry and stifle innovation.

5.1.3 Policy implications

It appears unlikely that there would be significant gains from enforcing standardisation in the water industry, primarily because of the low costs of converting. However further technical work is required to tell where the issues identified in Section 5.1.1 combine to support government intervention in standard setting, in particular in electricity.

Government interventions could include actions such as:

- mandating electricity frequency; and

- allowing industry to engage in co-operative standard setting, for example by making it clear that this is permitted under competition (anti-trust) laws.

5.2 Implications of free entry for economic regulation

This section focuses on regulation of the tariffs and service standards of infrastructure providers when free entry is allowed. Regulation designed to promote interconnection was dealt with in Section 5.1. Health, safety and environmental regulation were discussed in Section 4.5.

---

69 Kay, John, 1993, pp. 107-109.

70 For a sceptical view of network externalities, see Liebowitz, S.J & Margolis, S.E (1994) *Network Externality: An Uncommon Tragedy* Journal of Economic Perspectives Vol. 8, No. 2, Spring 1994 p. 133-150

In this section we first discuss briefly the general rationale for regulation. Then against this background, we analyse whether companies which enter an industry in response to liberalisation should be regulated.

5.2.1 Summary of rationale for utility regulation

Regulation is generally thought to increase efficiency. This view is usually based on the simple model of a monopolist producing a single good and selling it at one price to all who wish to buy it. This model, which many people learnt in school or first year university, shows that a profit maximising monopolist will restrict output and raise price above socially optimal levels.

More realistic models of infrastructure monopoly providers show that their profit maximising strategy will differ from that assumed by the simple model, and will entail lower efficiency losses. However, they will continue to make high profits. From this perspective, regulation may best be understood not as a mechanism to enhance efficiency, but as primarily concerned with achieving a socially acceptable distribution of income between producers and consumers of infrastructure services. On this view, the emphasis on efficiency in regulation is justified, but this emphasis is seen as being concerned with minimising efficiency loses which result from redistribution through regulation.

As Brooke Cowen and Cowen (1998) shows, unregulated providers of infrastructure services will often be efficient even if they are natural monopolies.[71] This is because:

- unregulated providers have the maximum incentive for productive efficiency; and
- unregulated providers are likely to charge tariffs which are more allocatively efficient than tariffs set by regulators. If an unregulated monopoly provider had perfect information, its pricing strategy would be to set tariffs which maximise total surplus, and then appropriate the surplus through a fixed charge exactly equal to each consumer's surplus. In reality, infrastructure providers will not be able to price-discriminate so perfectly. However they will generally have better information than a regulator, and stronger incentives to price in a way which maximises overall surplus.

In developing countries the efficiency arguments against regulation of infrastructure are bolstered by the fact that:

---

[71] Brook Cowen, Penelope and Cowen, Tyler, *Unregulated Privatization for Natural Monopolies,* Forthcoming in the Cato Journal.

- regulatory capacity is low, so regulation is likely to be poor; and

- the most common problem for utilities in developing countries is that they make inadequate profits, and therefore cannot finance investment. This means that regulation, which usually aims to limit monopoly profits, is a solution to a problem which does not exist in most developing countries.

Nevertheless, most countries regulate infrastructure providers. This is probably because people resent paying well above cost for very important or essential services. Politicians can tap this resentment by limiting prices, thus transferring rents from the provider to the consumer. This is often politically beneficial, since in many cases the provider has few votes, and consumers have many.

Arguably the history of electricity provision in the United States illustrates this process. Initially, power was supplied by private, unregulated companies. Increasing scale economies, and popular dissatisfaction with the prices and profits of these companies, led politicians to experiment with various ways of controlling utility prices, culminating in the current mix of Federal and State level regulatory commissions.[72]

Private investors welcome regulation in some cases. This is likely to be because they are aware of the political temptation to expropriate their profits after investments have been sunk. They believe that setting up in advance rules to determine the allocation of surplus reduces the risk of ad hoc expropriation.

In summary, it is generally most efficient not to regulate infrastructure, but political pressure to redistribute rents from producers to consumers means that most countries do regulate. Investors may welcome this because it provides certainty.

5.2.2 Economic Regulation of New Entrants

Bearing in mind the rationale for regulation in general, this section focuses on whether competitive entrants to an industry should be regulated.

In this discussion we assume that the incumbent is regulated initially.[73] If the incumbent were not regulated, there would probably be little reason to regulate entrants. We have seen that efficiency is likely to be enhanced by not

[72] Jacobson & Tarr (1996)

[73] By regulation we mean any kind of Government-imposed restriction on its freedom to set prices.

regulating, and if the political dynamic has not led to regulation of a monopoly incumbent, it is unlikely to lead to regulation of a new entrant.

If entry is sufficient to provide real competition in a market segment, then there is no need to regulate either the incumbent or the entrant. In the UK telecommunications sector for example, following privatisation in 1984, all of British Telecom's (BT) services were subject to a price cap. As competitive entry has occurred, regulation has decreased so that price regulations are only enforced over two specific aspects of BT's services.[74]

Often, however, the entrant and the incumbent will each enjoy some market power. For example:

- where the entrant serves a previously unserved area, there will be no competition. Each provider will have a local monopoly;
- where the entrant supplies an area previously served by the incumbent there will often be switching costs, limiting customers' ability to change between suppliers, and thus giving each supplier the ability to charge above cost; and
- even when switching costs are low, if there are only one or two entrants in an area, it is likely that the market will reach an equilibrium price which is above cost.

In all these situations, it is likely to be efficient to refrain from regulating the entrant. An unregulated entrant will have maximum incentives for productive efficiency. If the entrant is a local monopolist, it will have incentives to price in an allocatively efficient way. If the entrant faces some degree of competition from the regulated incumbent, its pricing will be constrained by the incumbent's prices. It seems unlikely that this would lead it to charge prices which were less allocatively efficient than the incumbent's, although this possibility would be worth researching further.

Most importantly, developing countries suffer from lack of investment in infrastructure. Regulating entrants is likely to reduce the profitability of entry, and therefore reduce investment.

Consider examples such as supplying sanitation services in an Indian slum, providing high quality water to an industrial estate in Trinidad, or establishing an internet service provider in Uganda. Each of these business

---

[74] The price regulations that remain are that all tariffs must be geographically uniform, and that the bills of the lowest 80% of residential customers must fall by 4.5% in real terms per year. While 80% appears high the spending patterns are very skewed so that the top 20% of residential customers contribute a disproportionately large amount of revenue.

ventures could be profitable. But each is highly risky, and could well fail. Taking into account the risk of failure, an investor would only invest if he believed there was also a chance of high profits. However a regulator will only see the businesses which succeed. After the business has succeeded, the risks will be smaller, and so the reasonable level of profits will seem lower. Knowing that a business will be regulated therefore cuts off much of the upside potential of an investment, while leaving the investor with the downside risk. Table 5-1 provides a numerical illustration of this problem

Table 5-1 : Economic Regulation can Deter Entry

		Expected return	
Scenarios	*Likelihood of each scenario*	*Unregulated*	*Regulated*
Failure	33%	0%	0%
Moderate success	33%	12%	12%
Best case	33%	36%	16%
Results			
Expected return		16%	9.33%
Investors' cost of capital		16%	16%
Investment decision		Invest	Do not invest

This table shows a hypothetical infrastructure investment. There are three possible outcomes of the investment: failure, moderate success, and best case. Each outcome is equally likely. There are two possible policy settings: unregulated or regulated. Under the regulated scenario, the provider will not be allowed to earn a return in excess of its cost of capital, which is assumed to be 16%.

The two right-hand columns show the returns the investor expects under each scenario. For the failure and moderate success scenarios, the returns are the same regardless of whether the business is regulated. This is because the returns are below the investor's cost of capital, so the regulator would not intervene. Under the best case scenario, the unregulated profit would be 36%. However in the regulated scenario, this would be limited to 16%.

In making the investment decision, the investor calculates the expected return, and compares it to his cost of capital. If the investor believes he will not be regulated, the expected return will be 16%, equal to the cost of capital, and the investor will invest. If the investor believes he will be regulated, the expected return drops to under 10%, and investor will not invest.

Note this problem will be most severe where the prospective entrant is a single project company. If the entrant has multiple projects in one regulatory jurisdiction, an arrangement could be worked out in which the entrant was

regulated to earn its cost of capital across a portfolio of projects, some of which would fail and some of which would succeed. Pharmaceutical companies in the UK are regulated on this basis, in the sense that they are allowed to earn a normal rate of return across their entire research and development portfolio.

There are three arguments in favour of regulating entrants which should be considered:

- achieving social and political objectives;
- encouraging investment by providing certainty; and
- providing equality of treatment between companies in the same industry.

Governments are often interested in redistributing rents from producers to consumers, even when this is not welfare maximising overall. This is particularly so when the producer is foreign owned. In this case, the rent distribution increases total national welfare even as it reduces global welfare. Where the Government has this objective, and where it is likely that entry will be highly profitable (and that entrants will be foreign owned), the Government may further its objectives by regulating the entrant.

Such cases are likely to be rare, however, as few free-entry infrastructure investments in developing countries will be so clearly profitable that regulation would not risk deterring entry. Where an entrant's investment is not clearly highly profitable, regulation may deter entry, as illustrated in Table 5-1. In this case, regulation may well reduce welfare for national consumers.

Investors know that if they are successful, there will be pressure to regulate them. For example, someone supplying electricity to an area for the first time may initially be able to charge a high price, as residents are grateful simply to get power. In time, however, residents will become accustomed to having electricity, and will start to compare the price with costs, and with prices in other areas. This will create dissatisfaction and lead to calls to regulate the local electricity company. In some cases investors might prefer to reduce this risk by establishing a regulatory framework in advance.

Our judgement is that providing entrants with certainty will seldom justify Government taking the initiative in regulating entrants. If investors want regulatory certainty, they will generally be able to suggest a suitable regime to the Government.

It may be seen as unfair to regulate the incumbent and not to regulate the entrant. Fairness between people is generally a policy objective, but fairness

between companies is seldom an objective in itself. Fair treatment of companies may be good practice, largely because it provides certainty and may provide competitive neutrality, thus promoting investment and efficiency.

In the case of free entry, the rationale for equal treatment of companies is greatly reduced. The incumbent, by assumption, has a near-monopoly position. Regulating the entrant in these circumstances will tend to deter entry, and thus reduce competition, investment and efficiency. The benefits of fairness will seldom outweigh these costs.

There are several countries which regulate the incumbent more closely than they regulate entrants. For example, Telecom New Zealand is subject to specific obligations with respect to local service and information disclosure. None of the other five telecommunications providers operating in New Zealand are subject to these obligations. Approaches similar to that adopted for telecommunications in New Zealand will generally be best practice.

In conclusion, it will generally be best practice not to impose economic regulation on new entrants. Exceptions to this are:

- where it is clear that entry will be highly profitable, and the Government values consumer welfare more highly than producer welfare;
- where it is clear that entry will be highly profitable, and that the entrant will be foreign owned; and
- where the entrant itself requests regulation.

5.3 Politics of monopoly and free entry

We have seen that:

- free entry will often be welfare enhancing; but
- most countries restrict entry.

The cause of this puzzling result can be found in part in the political dynamic which supports restrictions on entry. In the following sections we examine;

- the politics which create and sustain limitations on entry; and
- the political implications for a reformist government of allowing free entry.

5.3.1 Political dynamic of restrictions on entry

The politics of monopoly breaks into two parts:

- why are monopolies created; and
- once created, why do they persist.

Creation of legal monopolies

In some cases the genesis of legal monopolies in developing countries lies in the uncritical adoption of foreign ideas and models. For example:

- Trinidad, like many other former British colonies, took English utility law as the basis for its own post-colonial laws. Since the English law restricted entry, the Trinidadian law also restricted entry;
- in the 1970s, many people believed that the water sector was best structured into regional, vertically integrated monopolies with both service provision and resource management responsibilities. The previously diverse UK water industry was restructured along these lines in 1973. Many countries, including India and Brazil, restructured their industry in accordance with what was then regarded as best international practice; similarly
- former French colonies naturally looked to France as a model for their own utility sectors. Consequently, many countries established national electricity monopolies patterned on Electricite de France.

Socialist and nationalist ideology played an important role. In socialist countries, policy makers believed that planned state ownership was superior to private ownership and markets. Consequently, competitive entry was prohibited in most sectors, including infrastructure.

Some post-colonial countries were anxious to establish economic as well as legal independence. This was often interpreted to mean national ownership of key economic sectors, including infrastructure. As a result, infrastructure was often nationalised, or remained in government ownership after independence.

In these circumstances, free entry would have meant allowing competition with the government, which could be seen as harmful. It might also have meant allowing entry by foreigners in cases where the local private sector was insufficiently developed to finance and manage infrastructure projects. Again, this would have run counter to economic nationalist objectives.

There are cases in which short term or narrowly political objectives led Governments to grant monopolies which were contrary to the interests of the country over the longer term. In Jamaica, foreign exchange crises in the late 1980s and early 1990s led Governments to sell successive tranches of shares in Telecommunications of Jamaica (ToJ) to Cable and Wireless. The perceived riskiness of Jamaica as an investment destination, coupled with the urgency of the Government's need for foreign exchange, allowed Cable and Wireless to demand extensive monopoly powers. While the privatisation of ToJ brought initial gains in increased investments and improved quality of service, Jamaica's performance in telecommunications is now lagging again, and it is widely believed within the country that the monopoly privileges granted were too extensive and long-lived.

There may also be cases in which personal gain motivated grants of exclusivity. These tend not be documented. A possible example was the planned privatisation of the Argentine post office. A former Economy Minister, Mr Domingo Calvino, alleged that a bill to privatise the post office (which was unanimously approved by the Senate) was tailor made to give a postal monopoly to a politically connected businessman.[75]

Political dynamic sustaining monopolies

Legal monopolies are sustained by a variety of factors. We have seen that in some cases exclusivity will be justified by public policy objectives such as universal service and environmental protection. Even where these objectives could be better served through other mechanisms, it is perhaps not surprising that the economic theory of the 1940s to 1970s, which tended to emphasise economies of scale and a positive role for the State, continues to hold sway in many countries, both developed and developing. In other cases, the persistence of exclusivity may be nothing more than inertia; there may have been no call to revisit the assumptions which prevailed when the current laws were established.

In searching for an economic explanation for the persistence of monopolies in a variety of situations, including those in which exclusivity reduces overall welfare, a possible candidate presents itself in public choice theory.

Public choice theory (at least as expounded by economists such as Demsetz, Posner, Niskanen, Buchanen and Sowell) starts with Schumpeter's postulation that rent seeking is the primary objective of economic activity.[76] It applies this postulation to Government decisions, seeking to explain policy-

---

75 The Economist, May 30 – June 5th, p.35

76 Schumpeter, J.A. Capitalism, Socialism and Democracy, (1943), Unwin, ; and The Theory of Economic Development, Oxford University Press. (1961).

making in terms of the individual interests of officials, politicians, voters and others involved in policy decisions.

In the remainder of this section we use the tools of public choice theory to develop a possible explanation for the sustainability of restrictions on free entry. We provide empirical examples to demonstrate the plausibility of the theory. However, the emphasis on self-seeking behaviour which emerges is in part a reflection of the theoretical tool chosen. We do not mean to imply that this is only cause. Other factors, including both idealism and inertia are also important.

A public choice theory explanation of restrictions on entry starts with the observation that most monopolies generate rents, that is, returns in excess of those required to cover costs and remunerate investment. To the extent that the rents are due to government action, government may be able to influence distribution of the rent. The rent can then be distributed to political or personal advantage. For example:

- in France, holders of concessions for municipal services are sometimes alleged to pay off mayors or municipal officials, either directly in cash (a number of cases have come to court recently), or through political contributions, overseas trips, use of apartments in Paris, etc.;

- in several countries privatised utilities are an important source of campaign contributions; and

- it has been argued that one of the reasons for the nationalisation of the French electricity industry in 1946 was to end the undue and corrupting influence of utility owners in politics.[77]

In the above cases, rents are generated by natural monopoly factors, as well as by legal restrictions on entry. Government control over the rents often flows from ownership and other sources of influence as much as from control over entry. The various factors are generally impossible to isolate. Nevertheless, it could be assumed that:

- restrictions on entry generally allow additional rents to be generated;

- these rents will be partly under government control; and

---

[77] Jacobson, Charles D. and Tarr, Joel A. in Infrastructure Delivery: Private Initiative and the Public Good, IBRD (1996) pg.28, referencing Frost, R. L, Alternating Currents: Nationalised Power in France, 1946-1970, Ithaca, Cornell University Press. 1991.

- governments could benefit from control of these rents.

This can be so even when the entry restrictions reduce welfare overall. The reason is that the losers from the restrictions tend to be dispersed, while winners are concentrated.

For example, political patronage can be used to award party supporters jobs or contracts with monopoly utilities.[78] For the individual benefiting from patronage, the rewards are substantial. The inefficiency which results from such political appointment is eventually recovered in higher tariffs. However, because the tariff increase is spread across many customers, the cost to each one is small. Therefore it is not in any customer's individual interest to oppose the system, since the costs of lobbying for change would be high, while their personal benefit would be small.

There are also cases in which losers are more concentrated than winners, but the rent redistribution still makes political sense. For example, a populist Government might conclude that business people will support the opposition in any event. Therefore it will not lose any support by increasing utility rates to businesses. The revenue which this creates can then be used to subsidise rates to groups whose support might be swung toward the populist party.

It is also important to consider the dynamic within governments concerning rent distribution. If one treated the government as a single entity, one might conclude that the tax system would be the most effective way to collect and distribute rents. One would therefore question why the government would use less efficient systems, such as infrastructure licences. However, in many countries the tax system is controlled by the Minister of Finance. Other ministers may want their own source of funds and power. Therefore the minister in charge of an infrastructure area may have an incentive to create and maintain exclusive arrangements, even if this is not the most efficient way to generate rents from the point of view of the government as a whole.[79]

Perceptions and information costs have similar effects. The citizens in many countries are alert to the risk that governments may reduce overall welfare by redistributing rents for personal or narrowly political objectives. Citizens often resist these tendencies. Popular action is much more likely against a few gross or obvious abuses than against a myriad of small and complex schemes. For example, citizens are usually strongly opposed to tax increases. The cost imposed on citizens by inefficient infrastructure is usually much harder to see, and therefore inspires less opposition. In some cases this may provide an additional reason for governments to maintain restrictions on entry in infrastructure.

---

78 McCarthy (1994), Kerf & Smith (1996) Box 1.4 citing ESMAP (1993)

79 For a game-theoretic view of this dynamic, see Mauro (1995) p.705-706.

The factors which tend to promote political support for inefficient restrictions on entry also support government action to redistribute rents through many other mechanisms, including trade barriers, licensing in other industries, and industrial subsidies. Governments interested in reducing the effect of these factors therefore should consider both general and infrastructure specific solutions.

Various countries have introduced legal and constitutional measures designed at least in part to limit such rent redistribution. Trinidad, for example, will soon pass a new law governing utilities. It is likely that this will allow for the award of exclusive licences, but only where the Minister of Public Utilities certifies that this is necessary to achieve the objectives of the Act. Other countries require that all regulatory restrictions on business are subjected to a cost-benefit analysis before being passed. Competition law limiting the creation of monopolies could also help in some circumstances.

Conclusions on the Political Aspects of Creating and Maintaining Limitations on Entry

In conclusion:

- governments often restrict entry in infrastructure even when welfare would be increased by allowing entry;

- a range of factors explain initial restrictions on entry. Beside genuine attempts at good public policy, these include uncritical acceptance of external influences, and socialist or economic nationalist ideologies. In some cases, the restriction results from politicians or officials seeking personal gain or short term political objectives; and

- a similar range of factors explains the persistence of restrictions on entry. Public choice theory emphasises that restrictions on entry contribute to the stream of rents under government control. Even though such rents are inefficient, it is seldom in anyone's individual interest to change the system. This is because the losers from the system are dispersed, or politically unimportant. The winners from the system are sufficiently concentrated and politically important to maintain the system which benefits them. A political dynamic against free entry may be created.

This view of the political dynamic of restricted entry in infrastructure is analogous to public choice theories explaining other restrictions on competition, such as protectionism, and targeted industrial subsidies.

5.3.2 Political Implications of Allowing Free Entry

This section examines whether allowing free entry would assist a reforming government which is interested in increasing welfare overall.

Increasing welfare would include unravelling the rent redistribution mechanisms referred to in the previous section. For example, a reforming minister might want to:

- remove cross-subsidies in tariffs;
- reduce government subsidy of utilities; and
- reduce staff at an inefficient government-owned utility.

These reforms will be politically difficult precisely because they run counter to the political dynamic supporting rent redistribution. The losers from the reforms will often be more concentrated and politically influential than those who gain from it.

For example, reducing staff in a utility will be strongly opposed by workers at the utility. The winners will be consumers, who will face lower bills in the future. However, the change in the utility bill is likely to be small and uncertain (since it is spread over a wide group), so few consumers will provide political support for the reform. Politically, there would be little to counter-balance the opposition of workers.

In most countries, ministers could not avoid taking some responsibility for the losers in the system. However, the extent to which the minister is blamed may vary according to whether the losses are the result of direct ministerial action, or the result of decisions taken by independent parties acting at one remove from the minister.

For example, the minister may face a choice between:

- reforming a state electricity utility directly; or
- allowing free entry in generation and distribution of electricity, and allowing market pressures to take their course.

Assume the minister chooses the free entry approach. If the incumbent is inefficient, with a distorted tariff structure, cherry-picking entry to supply major industrial load centres is likely. As the incumbent loses industrial market share, it will be forced to drop industrial tariffs toward cost. Tariffs previously kept below cost through a cross-subsidy from industrial users will have to rise if financial viability is to be preserved. Shrinking revenues drive managers to cut costs by reducing staff. In the longer term, as fully commercial electricity entities become established and accepted, the

legitimacy of continued government subsidy to the former monopolist could be eroded.

In this scenario, the minister achieves all the reform objectives with only one decision; allowing free entry. Unpopular decisions to raise residential tariffs and reduce staff are taken by others, in response to independent, external competitive pressure. This could insulate the minister from the unpopularity of reform, while still allowing him to take credit for the vibrant, efficient and competitive electricity sector which will result.

Less fortunate outcomes are also possible. In some countries the minister would be not only blamed for the increase in domestic tariffs, which consumers will know he could have prevented, but also condemned for allowing foreign capitalists to exploit and recolonise the country through their cherry-picking behaviour. The incumbent utility would refuse to raise other tariffs or reduce labour, and instead would incur ever larger operating losses, which central government would have to fund.

These scenarios show that the usefulness of free entry in reforming a utility sector will be highly situation specific. In general, free entry will be most politically useful as a driver of reform where:

- there is a good prospect of competitive entry;
- competition is generally considered by the public or opinion formers to be desirable;
- there is acceptance of private ownership of important industries; and
- incumbent utilities are relatively autonomous and commercial.

Where few of these conditions hold, more hands-on reform will be required. Free entry could then be allowed toward the end of the reform process, after tariffs have been rebalanced, and the autonomy and commercial focus of existing utilities increased.

6 Likely significance of allowing free entry

In this section we briefly describe the kinds of changes that could be expected from allowing free entry in a sector in which it was previously forbidden

The effect of allowing free entry depends in large part on the situation at the time free entry is allowed. Where an existing monopoly provider is working well and meeting consumer demands at cost reflective tariffs, allowing free entry will have little or no immediate effect. Allowing free entry will have longer term benefits, however. It will provide a modicum of competitive pressure on the incumbent to maintain performance, and allow consumers an alternative in the event that performance does deteriorate. Perhaps most importantly, it ensures that if the incumbent does not take advantage of major technological changes in the future, entrants will be able to bring those innovations to the market.

Where an incumbent monopoly works well, but cross-subsidises some consumers and charges others above cost, 'cherry-picking' entry is likely. Cherry picking has been observed in the UK water market, and in international telephony in many countries. Entrants may or may not be more efficient than the incumbent. The likely effect is to force tariff re-balancing. This increases allocative efficiency. It may undermine social objectives, requiring the introduction of competitively neutral alternatives. Once tariffs are brought into line with the costs, the situation will be that of a well functioning monopoly charging cost-reflective prices.

Utilities which perform well are rare in developing countries, so it is most interesting to consider what happens when there is an inefficient incumbent failing to meet market demands, and free entry is allowed. While a wide variety of responses could be expected, some examples will give a flavour.

In Nigeria, Lee and Anas (1992) found that as a result of the unreliability of the monopoly provider (NEPA), 86% percent of all manufacturing firms own some form of generating capacity. The great majority of firms use NEPA as the primary source of power, but maintain sufficient back-up to power their entire operation in the event of a power failure. As a result, the private generating capacity has a load factor of only around 25%.

The authors surmise that were firms allowed to sell power to each other, they would respond in two ways:

- joint production – in which a large firm with existing facilities could arrange to supply neighbouring firms with back-up capacity; and

- 'shared production' in which firms club together to form an infrastructure pool – perhaps a separate firm or co-operative,

owned by the firms to be supplied, which will supply all firms in the club with infrastructure services.

The scope for economies of scale and increased load factors (effectively reduced reserve plant margins) mean that such arrangements could cut the total cost of electricity supply to firms significantly.

Similar responses could be expected in Uganda, for domestic as well as commercial and industrial customers. In Uganda self-generation of electricity is widespread. Many firms have standby petrol or diesel generators because of the unreliability of grid supply. These firms' generators seem to work for around 4-6 hours on average per day, which suggests a high level of un-utilised capacity.[80] Allowing free entry in Uganda for on-selling electricity would probably result initially in neighbours supplying power to each other. This could develop into club and co-operative arrangements, and eventually to the point that private companies set up specialising in the supply of electricity to areas currently unserved or poorly served.

In sanitation, Lee and Anas (1992) cite the private effluent collection and treatment facility on the Agbara Industrial Estate in Nigeria as an example of a utility pool. This facility was established by the private developer of the estate, and is operated by a management company. Similar arrangements can be envisaged in other industrial areas as environmental standards are tightened.

Many developing countries in the Caribbean and elsewhere depend on tourism. Where resorts line a beach, the costs of pollution can be largely internalised by a club of resorts in the area. Currently many hotels rely on package plants, which are non-core business, do not benefit from economies of scale, and often malfunction. If free entry were the norm, it is possible that resorts would club together to establish a single large sewage treatment plant, the management of which could be contracted out to a specialised private operator. In this case, it is probable that resorts would allow businesses and household in the area to connect to the system. A similar logic drove the establishment of a private water company supplying desalinated water to hotels and other customers in the Cayman Islands.

Free entry may also improve residential sanitation services in developing countries. The examples of the Orangi Pilot project and its replication in Lahore show that poor people in urban areas value sanitation highly, and are willing and able to pay for it, provided appropriate low cost technologies are used. However such willingness to pay extends only to removal of waste from the area. Additional Government intervention (such as providing for

[80] Ugandan Ministry of Natural Resources Energy Department, *Report on Rural Electrification* Prepared by Commissioner for Energy Ministry of Natural Resources (1997)

interconnection with an existing sewerage system) will be necessary to ensure proper treatment and disposal.

In water, the initial effects of free entry in many places will be to legalise current non-piped suppliers such as water carriers, water truckers, and people who sell to their neighbours. This may have beneficial effects of increasing certainty in the industry, allowing greater investment, and reducing the scope for racketeering.

One could expect that the next development could be for some people to specialise in on-selling piped water, initially through establishing private water vending kiosks. Low cost extensions of piped supplies into unserved areas could be the next development. Already squatter and slum settlements in many places are served with low-cost plastic piped networks taking water illegally from the mains. Legalising on-selling would provide the potential for people who are legally connected to establish similar low cost extensions to the network. The water company would benefit from the extra sales, and the greater security of a legal operation could, in time, lead to investment in higher quality piping, meters, etc. in the network serving the low income area.

Where water resources are available in the vicinity, entrants are likely to source water as well as distribute it. A small scale example is provided by the private vending of water treated by low cost UV disinfection in Manila, Philippines.[81] The aguateros of Paraguay show that such entry can be large scale, when the legal and physical environment permits. Aguateros operate wells to extract groundwater, may have as many as 800 customers, and lay substantial piped networks. Even in Paraguay, the legal environment for aguateros is uncertain. If free entry was unambiguously legal, it is likely that access to capital would be easier, and larger and more efficient networks could be constructed.

81 Solo & Snell (1998)

7 Conclusions

In conclusion, there should be a strong and general presumption in favour of allowing free entry in water and electricity industries. Where efficient incumbents supply most of the population at cost-reflective prices, allowing free entry will not often result in entry. This follows from the natural monopoly characteristics of water and electricity supply. Even in these cases, allowing free entry will seldom cause harm, and may do some good by pressuring the incumbent to remain efficient.

It is common in developing countries for the incumbent utility to be inefficient, provide a poor service to only a fraction of its potential customers, and to charge prices which bear little relation to costs. In these cases, the benefits of free entry can be considerable. Industries will benefit if high quality, reliable service can be provided by new entrants. Slums, rural areas and urban peripheries, which are not served by the incumbent, will be legally able to seek alternative sources of supply. There are numerous examples of communities and businesses having benefited from entry in these circumstances.

Free entry will also allow technological progress. The nature of technological change is that it is unpredictable. This makes it dangerous to limit entry in water and electricity just because it seems unfeasible currently. Future developments in fuel-cells, small-scale desalination, or other technologies as yet undreamed of may make these industries far more competitive in the future. The telecommunications industry shows the revolution which a combination of technological change and competition can bring to an industry once regarded as a natural monopoly.

In all these respects, water and electricity are little different from other sectors of the economy, in which free entry is the norm, and competition encouraged.

There are cases in which free entry will cause problems. Sometimes these problems will be so severe as to outweigh the positive effects, and may provide a rationale for restricted entry. The circumstances in which free entry might reasonably be restricted are summarised in Table 7-1

In reviewing Table 7-1 it must be borne in mind that its purpose is to show those cases in which the general presumption in favour of free entry may not apply. It shows cases in which allowing free entry might be harmful. In all cases, the harm from allowing free entry must be weighed against the damage which government imposed limitations on entry can do to efficiency, dynamism and transparency.

While governments will generally want to review each of the various factors listed in Table 7-1 in making decisions on whether to restrict entry, there is an

additional argument in favour of simply allowing entry in all or almost all cases. This has to do with the costs of complex government decision making.

As shown in Figure 3-1 earlier in the paper, policy conclusions may vary according to the administrative capacity of the government, and according to the nature of the problems faced.

Restricting entry places increased pressure on government or a regulator to address complex questions. As the regulatory decisions required become more complex, costs rise as a result of:

- increased costs of analysis and decision making within government, as each situation must be analysed afresh;
- increased costs on the private sector, as a result of delays and uncertainty; and
- increased losses from lobbying, corruption and other rent-seeking behaviour, as people try to manipulate the decision-making process to their own ends.

These costs will typically be greatest for countries with the least administrative capacity. It will often be the case that the costs of complex and discretionary decision-making outweigh the benefits of tailoring decisions precisely to each situation. In such cases simple rules – pure free entry or pure exclusivity – are best. Of the two, free entry will generally be the better rule.

Therefore, in the water and electricity sector of countries with limited governance capacity, a simple rule allowing free entry will often be the best policy.

Table 7-1 : Summary of recommendations

Administrative Capacity Issue	Very low	Moderate	Good
Inefficient entry and duplication Free entry could increase total production costs, as several firms race for the same market (chicken problem), or an entrant supplies only a subset of outputs with strong economies of scale ($_{1/3}$, $_{2/3}$ problem), or two providers target the same market niche, leaving others poorly served.	Allow free entry, since government capacity will be insufficient to identify and solve such problems.	Where there is an incumbent with significant sunk costs, allow free entry, and allow incumbent to set prices as it wishes. Where there is no incumbent, and the costs to entrants of establishing long term contracts prior to investing are high, consider organising a competition for the market. Period of exclusive franchise award should generally be limited to the period necessary for the entrant to become established, e.g. 5-7 years.	Where there is an incumbent with significant sunk costs, allow free entry, and allow incumbent to set prices as it wishes. Where there is no incumbent, and the costs to entrants of establishing long term contracts prior to investing are high, consider organising a competition for the market. Period of exclusive franchise award should generally be limited to the period necessary for the entrant to become established, e.g. 5-7 years.
Universal service Free entry may unravel a policy of cross-subsidising services to some groups by	Allow free entry. Where capacity is low, any scheme designed to promote universal service is likely to fail.	May be an argument for providing exclusivity in exchange for universal service obligations. However, this is	Allow free entry. Meet service objectives through competitively neutral mechanisms such as vouchers, or

Table 7-1 : Summary of recommendations

Administrative Capacity Issue	Very low	Moderate	Good
charging other groups above cost.	Better to allow the market to respond to consumers' needs.	only sustainable if the monopolist meets universal service targets. In many countries the monopolist fails to serve target groups, and free entry would be better.	industry-wide social obligation funds.
Exclusivity and private sector participation Free entry may decrease private sector willingness to invest.	Countries with low capacity will often be regarded as high risk. In addition, mechanisms to hedge risk may be hard to obtain. Private investors may demand exclusivity. There will often be a net gain from granting exclusivity in these circumstances. However, this should if possible be limited to the period the operator requires to establish itself – perhaps 5-7 years.	Requests for exclusivity should generally be resisted. In some cases, the benefits of granting exclusivity will outweigh the costs, but exclusivity should be for a limited period if possible.	Countries should try to avoid giving exclusivity. Investors concerns about risk can often be met through competitively neutral mechanisms, such as more sophisticated regulatory formulae. Transitional periods of exclusivity will be justified in some cases, for example to allow tariffs to be rebalanced and efficiency increased.
Time-tabling and routing problems In an integrated electricity system, some central	Not a problem in water – allow free entry. In electricity, allow free entry, subject to the provisions	Not a problem in water – allow free entry. In electricity, allow free entry, subject to the provisions	Not a problem in water - allow free entry. Allow free entry, subject to the provisions designed to promote

Table 7-1 : Summary of recommendations

Administrative Capacity Issue	Very low	Moderate	Good
mechanism may be needed to ensure efficient despatch.	designed to promote future interconnectability, outlined below.	designed to promote future interconnectability, outlined below.	future interconnectability, outlined below.
Environmental and Safety standards. Competitive pressure sharpens companies' incentives to cut corners with safety and the environment. At the same time, fragmentation of the industry makes enforcement more difficult.	Where there is no ability to enforce environmental and safety standards, generally allowing free entry will not make things significantly worse. Where there is some ability to enforce standards, and safety or environmental issues are of great importance, consider restricting entry.	Can provide a rationale for restricting entry. Need to bear in mind that: • provided personal liability law functions, this will often address safety concerns. May require operators to carry third party insurance. • environmental concerns in electricity can often be addressed just as well when there are multiple operators. Likely scale of entrants reduces grounds for restricting entry. • potential over-abstraction of aquifers or pollution of aquatic environments provides rationale for restricting entry in water	Allow free entry. Establish and enforce general, competitively neutral environmental and safety regulation.

Table 7-1 : Summary of recommendations

Administrative Capacity Issue	Very low	Moderate	Good
		and sanitation.	
Inter-connectability and technical standards Free entry may result in isolated systems developing along their own lines. It is likely that, in the future, interconnection would increase efficiency. However, incompatible technical standards may make this costly or impossible.	Allow free entry. If regulatory capacity is sufficient, consider encouraging or requiring compliance with basic standards in electricity, such as frequency. Allow industry standard setting and co-ordination.	Allow free entry. Consider encouraging or requiring all entrants to adopt the same basic standard in some areas, e.g. electrical frequency. Allow industry standard setting and co-ordination.	Allow free entry. Consider encouraging or requiring all entrants to adopt the same basic standard in some areas, e.g. electrical frequency. Allow industry standard setting and co-ordination.
Economic regulation	Do not regulate entrants unless they request it.	If entrant likely to remain small compared to a regulated incumbent, or if market is becoming increasingly competitive, do not regulate. In cases of significant entry into a new area which is likely to remain a monopoly, consider regulation through franchise-bidding.	If entrant likely to remain small compared to a regulated incumbent, or if market is becoming increasingly competitive, do not regulate. In cases of significant entry into a new area which is likely to remain a monopoly, consider regulation through franchise-bidding.

Annex 1: Bibliography

Barron's online; *Crowed Skies: Will satellite phone systems come crashing down on investors' heads,* June 15 1998

Brook Cowen, Penelope and Cowen, Tyler (1998); 'Deregulated private water supply: A policy option for Developing Countries', *The Cato Journal, Vol.18, No. 1*

Chernow, R (1997); *The Death of the Banker: The decline and fall of the great financial dynasties and the triumph of the small investor,* Vintage Books, Random House Inc., New York

Commerce Commission (1996); *Report on the proposed acquisition of shares in Ansett Australia by Air New Zealand*

Crandall and Ellig (1997); *Economic Deregulation and Customer Choice: Lessons for the Electric Industry*, The Centre of Market Processes Inc.

Crane, Randall (1994); article in *Water and Development announced in Voices from the City, the Peri-urban network newsletter,* April 1994; cited in Solo and Snell (1998)

de Boles, David and Evans, Lewis (1995); *The economic efficiency of telecommunications in a deregulated market: The case of New Zealand,* Graduate School of Business Government and Management Wellington University Working Paper 8/95

De Soto, Hernando (1989); *The Other Path,* New York, New York, Harper and Row; cited in Mauro (1995)

Drangaert, Pili, Malgarejo, Silvio, Kemper, Karin and Bakalian Alex (1997); *Water Supply in Paraguay: The Role of Small Entrepreneurs*, World Bank

Economist, The, 12 August 1995; *Investing in Africa: A new scramble*

Economist, The, May 30 – June 5th 1998 p.35

ESMAP (1991); *Republic of Yemen Household Energy Strategy Study, Phase 1 A Preliminary Study of Northern Governorates*, Report No. 126/91 Energy Sector Management Assistance Program (ESMAP)

ESMAP (1993); *Nigeria: Issues and Options in the Energy Sector*, (No.11672-UNI)

ESMAP (1997); *Institutional Development for OFFGRID Electrification in Laos (Draft),* UNDP and World Bank Energy Sector Management Assistance Programme, March 1998

Financial Times Business Report: Technology (Q3:35) 13 February 1997

Frost, R. L (1991); *Alternating Currents: Nationalised Power in France, 1946-1970*, Ithaca, Cornell University Press,

Irwin, Tim (1997); 'Price Structures, Cross–subsidies and Competition in Infrastructure', *Public Policy for the Private Sector*

Irwin, Tim; *Free Entry Competition in Infrastructure as a Response to Poor Governance,* (Preliminary Draft)

Jacobson, Charles D. and Tarr, Joel A (1996); Private Initiative and the Public Good, *Infrastructure Delivery*, IBRD

Kay, John (1993); *Foundations of Corporate Success,* Oxford University Press

Kerf, Michel and Smith, Warwick (1996); *Privatising Africa's Infrastructure: Promise and Challenge*, IBRD World Bank

Klein, Daniel and Moore, Adrian; 'Schedule jockeying and route swamping: bus markets in Britain need kerb rights', *Institute of Economic Affairs* (June 1997)

Klein, Michael (1996); *Competition in Network Industries,* Policy Research Working Paper #1591, World Bank

Krueger, Anne (1993); 'Virtuous and Vicious Circles in Economic Development', *Papers and proceedings of the American Economics Association* LXXXIII (1993); cited in Mauro (1995)

Lee, Kyu Sik and Anas, Alex (1992); 'Costs of Deficient Infrastructure: The Case of Nigerian Manufacturing', *Urban Studies,* Vol. 29, 1992

Lewis, Maureen and Miller, Ted R. (1997); 'Public - Private Partnership in Water Supply and Sanitation in Sub-Saharan Africa', *Health Policy and Planning*, (2) 1

Liebowitz, S.J & Margolis, S.E (1994); 'Network Externality: An Uncommon Tragedy', *Journal of Economic Perspectives* Vol. 8, No. 2, Spring 1994 p. 133-150

Mauro, P (1995); 'Corruption and Growth', *Quarterly Journal of Economics,* August 1995 p.681-712

McCarthy, Stephen (1994); *Africa: The Challenge of Transformation,* St Martin's Press, New York, New York

Milgrom and Roberts (1992); *Economics Organisation and Management*

Negroponte; 'RJ- 11', *Wired*, Issue 6.04 April 1998.

OECD Pricing of Water Services, (1987)

Reuter News Service – Far East Reuter Textline (Q2:60) 31 March 1998

Schumpeter, J.A. (1942); *Capitalism, Socialism and Democracy*, New York, Harper

Schumpeter, J.A. (1961); *The Theory of Economic Development*, Oxford University Press

Solo, Tova Maria & Snell, Suzanne (May 1998); *Water and Sanitation Services for the Urban Poor – Small Scale Providers: Profiles and Typology*, UNDP-World Bank Water and Sanitation Program

Tirole, J; *The Theory of Industrial Organisation*

Tynan, Nicola and Prof. Cowen, Tyler (1988); *The Private Provision of Water in 18th and 19th Century London*, First Draft, May 1998, Department of Economics, George Mason University. (Note – this paper is not to be cited without the author's permission. Since the recipients of this report already have a copy of the paper, we have cited it. However, permission should be sought before this report is more widely circulated)

Trinidad Sunday Guardian 7 June 1998 page 3

Uganda Ministry of Natural Resources Energy Department; *Report on Rural Electrification*, Prepared by Commissioner for Energy Ministry of Natural Resources (1997)

Wasserstein, Bruce (1998); *Big Deal: The Battle for Control of America's Leading Corporations* Warner Books Inc.

Whittington, Lauria and Mu (1989); *Paying for urban services: A study of water vending and willingness to pay for water in Nitsha, Nigeria* Infrastructure and Urban Development Department, Case Study, Report INU 40; cited in Lee and Annas (1992) p.1088

Footnotes from Table 2-1: Spectrum of entry policies

[i] Tynan, Nicola and Prof. Cowen, Tyler (1998) *The Private Provision of Water in 18th and 19th Century London*, First Draft, May 1998, Department of Economics, George Mason University

[ii] Solo, Tova Maria & Snell, Suzanne (May 1998) *Water and Sanitation Services for the Urban Poor – Small Scale Providers: Profiles and Typology* UNDP-World Bank Water and Sanitation Program, page 9, citing Randall Crane article in Water and Development announced in Voices from the City, the Peri-urban network newsletter, April 1994

[iii] Solo & Snell, (1998)

[iv] Economist 12 August 1995 *Investing in Africa: A new scramble*

[v] Whittington, Lauria and Mu (1989) *Paying for urban services: A study of water vending and willingness to pay for water in Nitsha, Nigeria* Infrastructure and Urban Development Department, Case Study, Report INU 4o; cited in Lee and Annas (1992) p.1088

[vi] Drangaert,Pili, Malgarejo, Silvio, Kemper, Karin, and Bakalian Alex, *Water Supply in Paraguay: The Role of Small Entrepreneurs* World Bank (1997)

[vii] Solo & Snell (1998)

[viii] Solo & Snell (1998)

[ix] Solo & Snell (1998) p 9

[x] Solo & Snell (1998)

[xi] Solo & Snell (1998)

[xii] Solo & Snell (1988)

[xiii] London Economics research

[xiv] Lee, Kyu Sik and Anas, Alex (1992) "Costs of Deficient Infrastructure: The Case of Nigerian Manufacturing" *Urban Studies, Vol. 29*

[xv] London Economics research

[xvi] Lee & Anas (1992)

[xvii] Lee and Anas (1992) p.1088

[xviii] Economist 12 August (1995)

[xix] ESMAP (1991)

[xx] ESMAP (1998)

[xxi] London Economics research

[xxii] London Economics research

[xxiii] London Economics research

[xxiv] London Economics research

[xxv] London Economics research

[xxvi] London Economics research

[xxvii] Lee & Anas (1992)

[xxviii] Chernow Ron (1997) *The Death of the Banker: The decline and fall of the great financial dynasties and the triumph of the small investor,* Vintage Books, Random House Inc, New York

[xxix] London Economics research

[xxx] New Zealand requires telecommunications and other infrastructure providers to register with the Ministry of Commerce. Registration is automatic once a small number of competitively neutral conditions (such as payment of fee to cover administrative costs and information reporting) are complied with. London Economics research

[xxxi] Barron's online *Crowed Skies: Will satellite phone systems come crashing down on investors' heads* June 15 1998