Infrastructure Investment Under Uncertainty[1]

Lewis Evans

School of Economics and Finance, and

NZ Institute for the Study of Competition and Regulation

Victoria University of Wellington, New Zealand

1. Introduction

Many volatile factors influence the performance of infrastructure and these yield a range of uncertainties when forward-looking investment decisions are being considered. This paper is restricted to consideration of physical infrastructure which has a wide spectrum of such factors. It includes physical events such as earthquakes that are beyond the influence of human-kind, other events for which there is a very small probability of occurrence and events that will almost certainly occur at some point within any reasonable period of time. It also includes economic events relating to uncommon financial episodes and common, but uncertain, volatility in demand and cost. Rare physical events have implications for investment in infrastructure that provides some mitigation of the effects of these events. In so doing, there is a trade-off between providing in advance for remotely likely but substantial events in specific, and usually costly, redundancy infrastructure, and having an economy with the resources to deal ex post with natural disasters. Obviously, some intermediate position will be socially desirable.

This paper considers investment in infrastructure taking into account more immediate risks. It argues that demand should be responsive to infrastructure's direct and indirect costs and risks; and that, where economically feasible, pricing[2] will facilitate management of these risks and so enable a

---

[1] This paper draws heavily on work with Graeme Guthrie.

[2] Indirect costs include costs imposed by individuals that affect others. These suggest prices such as congestion prices that enable consumers of infrastructure to express their demand for it while paying the cost of externalities induced by their use of the infrastructure.

desirable level of investment in infrastructure. Much infrastructure - e.g. roads, electricity and gas transmission, broadband and telecommunications networks - provide platforms on which consumers interact in various ways that affect the utilisation of the platform. Without consumers revealing their willingness to pay for these platforms, investment is unlikely to meet the test of being socially desirable. This issue is placed in perspective below by consideration of the effect of incentive regulation on investment.

Infrastructure investment once made is sunk - i.e. not recoverable in nearly its entirety - and typically entails economies of scale in investment: even in infrastructure maintenance expenditure.[3] These features and uncertainty in demand mean that provision of infrastructure is investment in capacity that services demand rather than in demand *per se*. When combined with volatility, these features complicate the evaluation of infrastructure investment.

2. Volatility and Economies of Scale

There is volatility in both demand and cost, with the extent of volatility depending on the nature of the industry. Technological change affects cost and demand, and where it is rapid - as in telecommunications - its effects on investment decisions can be significant.

Demand volatility, and hence risk, arises where there is competition in modes of delivery - e.g. as between road and rail, and for gas pipelines as between alternative fuels and locations of customers. Even the utilisation of gas pipelines can exhibit very high volatility in demand at different locations in the network of pipes. To illustrate; between 1995 and 2003 the flow through one of Natural Gas Corporation's piplines fell from the capacity of 500 Terajoules (TJ) per period to 50 and back to 400 TJ: during this period other pipes were stranded as their gas flows fell to zero. This substantial variation evidences very considerable demand risk that must be assssed in advance when evaluating investment in capacity. that will be irreversibly sunk once in place.

Cost uncertainty also arises due to variation in technological change, and a range of other factors. PBA (2004) report that cost variation can be attributed to: the price of inputs such as labour and materials, the level of competition, the level of supply and demand, project size and location, legal

---

[3] Economies of scale in investment mean that the larger the quantum of investment the lower the cost per unit of service or output of the additional capacity.

and regulatory requirements, constraints imposed by local authorities, choices between new construction versus use of established locations, design and construction standards, and the efficiency of the project and contract management. While cost uncertainty is reduced as a project becomes more specific - e.g. in location and design - much uncertainty may remain. An analysis of tenders for thirty roading projects in Auckland, Christchurch and Wellington as reported by Transit NZ (2006) suggests that on average the range of tenders for the same project was 26% of the maximum tender.

PricewaterhouseCoopers (2004) report on project quotes for four categories of investments across six electricity lines companies. The results reported in Table 1 indicate a very substantial variation in potential construction costs.[4] By way of illustration, if the quotes were normally distributed, a lines company contemplating an urban 11kV project would be of the order of 95% certain that the spread of quotes would be 55% - 155% of the average quote received. Variation, and thus prospective risk, is reduced by negotiation as the project is finalised but risk remains.[5]

Table 1: Variability of Infrastructure Construction Costs

	Underground	Transformer	11kV Urban
Coefficient of Variation	17.8%	40.1%	27.8%

Economies of scale in investment arise where the larger the capacity provided by the investment the lower the per-unit cost of the extra capacity. This is illustrated in Figure 1, where economies of scale are 10% - i.e. 0.9 units of investment are required to produce 1 unit of capacity: constant economies of scale in investment would be where investment was 1 for 1 with capacity. In Figure 1, investment in two steps obviously has a much higher investment cost than investment in a single step.

The conjunction of volatility and investment economies of scale complicates infrastructure investment decisions. On the one hand, a large increment in capacity will yield lower construction costs per unit of capacity than will a multi-stage investment. On the other hand, with uncertain

---

[4] The Coefficient of Variation is the Standard Deviation of the quotes for the same project divided by the average quote for that project.
[5] The risk may well be shared between the investor and the construction company.

demand growth, there may be inadequate demand for the larger capacity. Typically, capacity is expanded iteratively trading-off these two factors: where demand is more uncertain, the higher is the likelihood of the smaller increment in capacity being socially desirable; despite its higher cost.

Figure 1: Economies of Scale in Investment
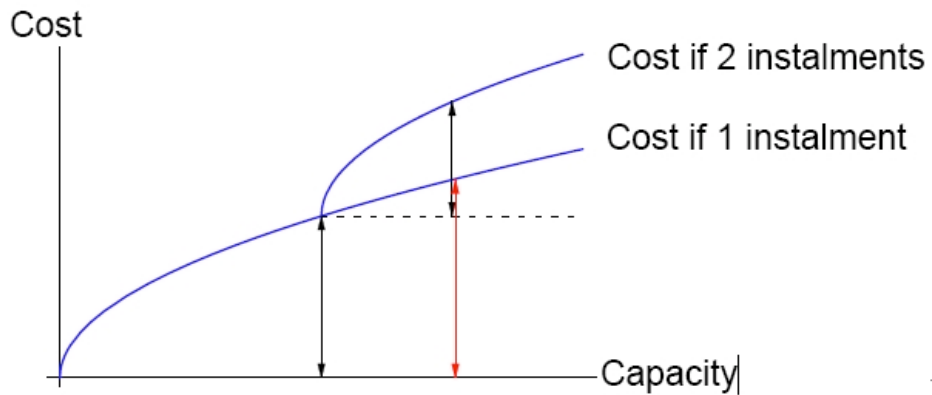
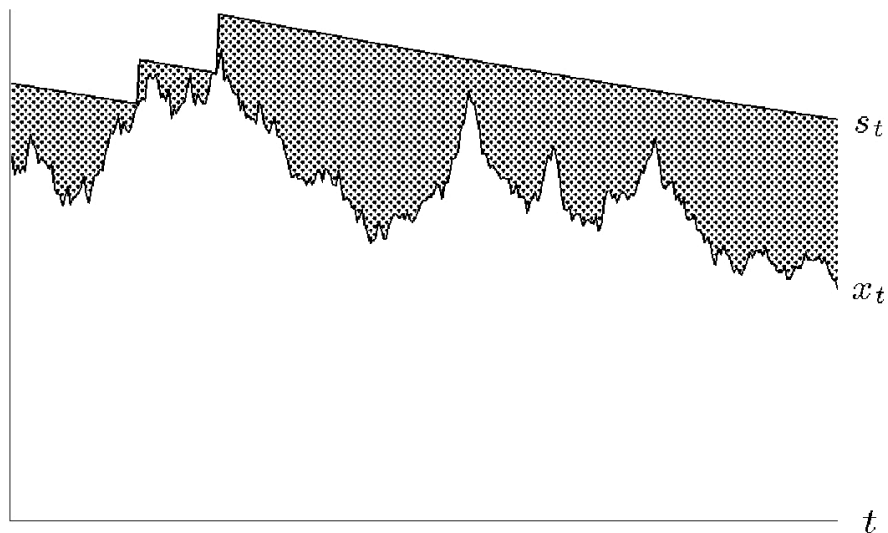**Figure 2: Demand, Capacity and Scale Economies**

Figure 2 indicates the decision rule in the case of volatile demand, and 10% economies of scale in infrastructure investment. Demand and capacity are on the vertical axis and time on the horizontal.

Demand (x) is volatile and must be served, and capacity (s) is irreversible (sunk) but declines without investment at a fixed rate of depreciation. The socially optimal decision rule is to invest whenever demand equals capacity and at that time increase capacity beyond the amount required to meet immediate demand. This decision rule is a consequence of the presence of investment economies of scale (see Evans and Guthrie (2006)), and it is affected by the variability in demand.[6] Building an extra unit of excess capacity allows the firm to connect new customers in the future without investing (at higher cost), but it destroys the option to wait and assess if such customers will arrive.

3. Project Evaluation and Regulation

The concurrence of risk and irreversible investment materially affect investment decision-making (see Dixit and Pindyck (1994), and Guthrie (2009)). The key effect is to render it socially desirable[7] that the variation in demand and cost be a critical element in the investment decision. In particular, investments that seek to maximise the expected present value of the sum of producers' and consumers' surpluses into the foreseeable future should consider the timing of the investment, not just whether, if carried out, it will be socially beneficial at the date of evaluation. In situations of risk and irreversible investment it is generally desirable that there is some delay beyond this date. The delay enables some resolution of uncertainty. If the investment climate improves, much less is lost by delay than would be lost by immediately (irreversibly) investing and the investment turning out to be bad because demand (cost) turns out to be low (high). The larger the risk - or variation of demand and cost - the larger the private and social benefit of the option to delay. Economies of scale may induce a longer waiting period to invest because increased surety of demand increases the sense of building a larger expansion in capacity and thereby gain the cost advantages of economies of scale in investment.

The interactions among risk, irreversible investment and economies of scale has been the Achilles heel of incentive regulation of infrastructure assets. It is useful to consider why this is so because it foretells the sorts of institutional arrangements that facilitate socially desirable investment in infrastructure. In New Zealand and in some other countries it was proposed that such infrastructures as transmission, pipelines and telecommunications be subject to incentive regulation

---

[6] And variability in cost, where this exists.
[7] Although some firms' decisions may differ from those preferred by society.

in which the regulated price be set at a level that financially just supported the most efficient firm in its delivery of services, independently of the actions of the firm being regulated.[8] The efficient price to be calculated was set as a price that would just enable a hypothetical, efficient firm to exist and provide existing services. The effect of this on firms' decision-making is illustrated by examining its effect on the valuation of the infrastructure firm.

A firm looking forward from some date t has a valuation given by:

Value(t) = Expected Present Value of Revenue less Expected Present Value of Costs

The expected present value of costs contains the sunk cost of the capacity in existence at date t, as well as expected future investment in the network. Consider the effect of this incentive regulation price setting where demand has to be served, there is 10% investment economies of scale and uncertainty about future costs and demand: both sources of uncertainty are reflected in the valuation of the firm that owns the infrastructure. The valuation makes some allowance for economic uncertainty (systemic risk) in the level of its discount rate but it does not include uncertainty about the very many other risks to the project's social and private viability. In this setting, Evans and Guthrie (2006) depict a firm that holds an *existing* capacity of 100 units and an associated rate base of K(t), and a regulator setting allowed revenue for the infrastructure provider as follows:

Case I: just sufficient revenue for the firm to keep operating but not enough to start up: this requires setting revenue to cover the expected cost of additional investment but it disallows accumulated past investment;

Case II: just sufficient revenue for the firm to start-up and keep operating;

Case III: just sufficient revenue for the firm to start-up, keep operating and not lose value when it expands capacity.

The Case I firm is just willing to operate using its existing assets, that is, those put in place in the past and depreciated. Because it is earning no return on its existing assets the revenue it receives just covers its expected capacity expansion cost. At low demand it makes little profit and hence has a low

---

[8] In a number of countries this approach has been applied to calculating access prices for telecommunications services - see for example, the widely used forward looking cost concept of total service long run incremental cost (TSLRIC) . In New Zealand, this regulation was proposed for lines companies by the Commerce Commission but was never actually implemented.

valuation at that level of demand. But its profit increases as demand increases – and its network is more fully utilised - until the point where the firm's anticipation of the cost of investing in expanded capacity outweighs the revenues per unit of demand. As demand approaches capacity, the probability of having to invest in expanded capacity increases to the point that the expected cost outweighs the revenue allowed per unit of demand. Thus the value of the firm declines: by enabling the firm to just cover expected investment cost, the value of the firm where demand equals capacity is zero. The Case I firm would never start up for its value lies below its replacement cost: this situation arises where existing assets are not allowed to, or cannot, earn a competitive rate of return. The decline in value at higher levels of demand means that the firm is contemplating investment in capacity that will have a negative pay-off to it.

In Case II, the firm is allowed just enough revenue for it to start-up and continue operating. Its situation is as for Case I, but with a minimal revenue stream covering both existing assets and additional, but prospective, investment. This firm will have a valuation greater than its replacement cost at moderate levels of demand, but it will try to avoid investment in additional capacity, because as demand approaches capacity the firm's valuation falls, even below its optimised replacement cost (ORC). The revenue assigned this firm is insufficient for it to invest and maintain its value when it faces network expansion. The reason for this result is that revenue will be reset as the revenue required to just support a hypothetically efficient firm that produces the same level of services as the firm in Case II. This revenue will be based upon the cost of building a *single* network and hence must be lower than that required to just support a firm that makes *incremental* decisions over time because of the presence of economies of scale. Put another way, while the firm makes incremental capacity decisions anticipating uncertain demand, the regulator sets the price after the firm's decisions applying the economies of scale to the whole firm and with no uncertainty about demand. The *ex post* actions of the regulator take place with more information than the firm had when it made decisions and it utilises economies of scale more extensively.

If there were no economies of scale; but rather constant returns to scale the firm does not lose value by expanding network capacity and thus has the incentive to invest in new capacity as required under incentive regulation. Comparison of consstant returns to scale and Case II illustrates why incentive regulation fails in the case of economies of scale in investment. Scale economies must produce a conflict between the regulator and the firm in which the firm seeks to reduce its

investment relative to that desired by the regulator.

In Case III, the firm has sufficient revenue that its valuation does not decline as demand approaches capacity. Evans and Guthrie, *op cit*, explain that, in the presence of investment economies of scale, this desirable state can only be achieved if the firm is allowed an inordinately large return on its assets: a return that would not be contemplated by a regulator. It is for this reason that pure incentive regulation fails where there are economies of scale in investment. These economies exist for most infrastructure and hence pure incentive regulation is unsuitable for it. In much price regulation oure incentive regulation has been replaced by historical cost regulation where there is intense regulatory oversight and approved infrastructure investment projects are included as capital in the rate base.

4.   Demand And Investment

Cases II and III illustrate that where demand must be satisfied at prices that approach the cost of infrastructure services, it will be a challenge to achieve the socially desirable level of investment where there are investment economies of scale. If price is set at a level that just covers the cost of a replacement firm, society will have to subsidise the infrastructure provider to achieve the desirable level of investment.[9] If it sets a price that just covers the incremental costs the firm incurs with its sequence of investments so that the firm is agreeable to investing, it will no longer be incentive regulation: it will be approved investment management. In this situation demand management becomes as important as investment management. In Case II, the firm's conflict with the regulator might be resolved by allowing excess demand to reach some level before investment takes place, even in the presence of investment scale economies. Indeed, this has been an approach long advocated by some.[10] The income generated by the jump in number of customers using the infrastructure at the time of investment enables the firm to not lose value at the time it invests. Whether, this means that the firm invests at the socially desirable time will be affected by whether it has competition or is subjected to regulation that precludes it making excessive rents from

---

[9] This is the dynamic analogue to the classic static depiction of natural monopoly. The need for a subsidy changes the concept of the desirable level of investment.

[10] The argument was advanced as long ago as 1970 by Baumol and Bradford in a setting without risk but with growing demand. A second approach not considered here is to charge bundled, or two-part tariffs: these may reduce consumer surplus at any point in time but bring forward investment in capacity to the benefit of future consumer and producer welfare.

congestion.

Excess demand requires prioritisation of use of the capacity, and this may be achieved by pricing where it is economic, or by congestion broadly conceived.[11] Congestion pricing for infrastructure importantly allocates the capacity to those that most value its use, and it provides information about the willingness to pay for an expansion in infrastructure. Both features are highly desirable if not essential for investment in infrastructure is to be at a socially desirable level.

References

Baumol, William A. and David F. Bradford. 1970. "Optimal Departures From Marginal Cost Pricing", *American Economic Review*, 60(3), 265-283.

Dixit, Avinash K. and Robert S. Pindyck. 1993. *Investment Under Uncertainty*, Princeton University Press, 468pp.

Evans, Lewis and Graeme Guthrie. 2006. "Incentive Regulation of Prices When Costs are Sunk", *Journal of Regulatory Economics*, 29(3), 239-264.

Guthrie, Graeme. 2009. *Real Options in Theory and Practice*, Oxford University Press, 414pp.

PricewaterhouseCoopers (PWC). 2005. *Cross Submission Intention to Declare Control on Unison Networks Limited*, Draft Report, 15 December 2005.

Parsons Brinkerhoff Associates (PBA). 2004. *Review of Optimised Deprival Valuation Handbook, Replacement Costs of System Fixed Assets*, Prepared for Commerce Commission, 31 August 2004.

---

[11] Congestion can take various forms that represent reduced service quality - e.g. delays and poorer service - and be managed by prices, administrative rules or laissez faire which is unlikely to be socially desirable for infrastructure.

Transit New Zealand. 2006. *www.transit.govt.nz*, accessed September 2006.