NETWORK SEPARATION AND INVESTMENT INCENTIVES IN TELECOMMUNICATIONS

August 2007

Martin Cave and Chris Doyle
Warwick Business School

University of Warwick
Coventry, United Kingdom

Table of Contents

	Executive summary	3
	Introduction	5
1.	Separation in industry 1.1 Separation and integration in sectors of the economy 1.2 For what problems is separation a remedy? 1.3 Separation options 1.4 OECD recommendation on separation in regulated industries	7
2.	Separation in the telecommunications context 2.1 Where to separate 2.2 Operational separation 2.3 Structural separation 2.4 Investment under different types of separation 2.5 An interim summary	12
3.	Economic theory on separation vs. integration	20
4.	Case Studies A. Regulated sectors 4.1 Investment incentives under price regulation 4.2 Railway sector 4.3 Airports 4.4 Energy 4.5 Separation in regulated industries: an overview B. Non-regulated sectors 4.6 Vertical integration in the US automobile industry 4.7 Personal computer and gaming industry 4.8 Contracting in the energy sector	23
5.	Implications for telecommunications	34

[This paper was prepared with financial support from eircom.
Address for correspondence: Martin Cave, Warwick Business School, Coventry CV47AL, UK. Martin.Cave@wbs.ac.uk]

Executive summary

- If a sector's vertical structure contains segments, which exhibit both monopoly and more competitive activities, there is scope for the vertical transmission of market power. A vertically integrated firm has the means and motive to practise non-price discrimination in the activity where it is dominant to weaken or eliminate competitors in potentially competitive activities. Recognition of the prevalence of this phenomenon has triggered the current wave of interest in separation.

- The fixed telecommunications sector, for example, is the subject of much discussion in relation to separation. This is largely a result of the Office of Communications' acceptance of BT's undertaking in 2005 to implement the operational separation of its access network, or "LoopCo", business in the United Kingdom, with the exception of Northern Ireland.

- A standard objection to the imposition of structural separation in telecommunications is that it will, *inter alia*, impede the desirable co-ordination of retail, core network and access network investment activities. In fact, the same charge can be levied against functional separation which, if it is to be implemented in a way which prevents discrimination, requires the 'ring-fencing' of both operational and investment activities undertaken by the separated entities, in order to exclude the possibility of abusive strategies.

- Public policy analysis – and, in particular, a 2001 recommendation from the Organisation for Economic Co-operation and Development (OECD) -- calls for an appraisal of alternative structures of a firm on their merits. The key points to emerge are:
  - Separation of vertically integrated firms and sectors has a clear benefit: the elimination of the potential for anti-competitive behaviour.
  - Separation may also have costs (loss of economies of scope, co-ordination costs, etc.); the benefit and costs are likely to vary from sector to sector.
  - Functional separation is a comparative novelty in the telecommunications sector, and experience of structural separation is even more limited.
  - In relation to the co-ordination of investment, the two types of separation appear in practice to impose similar, if not identical challenges.
  - From the standpoint of achieving pro-competitive benefits of separation, the optimal point of separation depends on the scope for competition in the market in question.
  - As far as investment co-ordination is concerned, the difference between operational and structural separation diminishes on inspection; it is likely that objections to the one form of separation will apply to the other.

- The theoretical literature on separation and integration also points to possible problems, which might arise in a separated structure if one party behaves opportunistically. A typical example of such behaviour might be for one party to wait until a partner has purchased a specialised asset, and then seek to lower the price paid for the output of such an asset, in the knowledge that its owner has no alternative outlet for its production. Vertical integration would solve this problem. But this example hinges upon the parties' inability to devise a contract (for example, a long-term one) which would overcome the problems of such post-investment opportunism.

- The theoretical arguments against structural separation assume that contracts, which will deter opportunistic behaviour once investments have been made, are very costly to write or cannot be written, and that regulation will be unable to prevent the exercise of market

power. Separation thus will lead to investment coordination problems and probable vertical transmission of market power. These assertions, which are characteristic of much of the theoretical literature, are highly contentious and deserve testing in practice.

- A further theoretical objection to separation is that if both stages of production exhibit market power, then the supplier of network services will take an excess profit, which will then be added to when the retailer imposes its excessive margin. But regulators seek to prevent precisely such exercises of market power by price control. Thus theoretical arguments against separation are based on pessimistic assumptions about both contracting and regulation.

- To see whether these concerns about structural separation are justified, the authors examine the advantage of separation and integration in a number of illustrative case studies of both regulated and unregulated sectors. The high-level findings are:
  - Judgements about vertical separation in railways are coloured by the performance of the flawed structure adopted in the UK. Across a broader sample of countries, opinion is more divided.
  - Problems characteristic of all regulated sectors apply to finding optimal investment incentives for airports, which reflect customers' needs for expansion. These have led to a number of palliative remedies, including constructive engagement and investment triggers.
  - Energy is an area where a vertical structure causes clear detriments to consumers, often delivered by co-ordinated under-investment.
  - More recent empirical research in non-regulated sectors effectively debunks a mistaken but widely cited argument against the effectiveness of contracts and in favour of integration.
  - A key learning from the personal computer and gaming industry is that firms can manage vertical issues in a highly dynamic complex industry in a sophisticated way, which neutralises the fear of opportunistic behaviour across a transaction boundary;
  - The energy asset field provides examples of contracts can be devised which do take the strain, and that investment is co-ordinated subject to them across contractual boundaries. The extensive Economics literature on contracting, which covers many sectors, contains numerous other examples.

- By definition a separated structure imposes heavier demands on contracting. But the evidence of academic research cited and case studies presented is that contracting can, in most cases, take the strain, by such means as long term or risk-sharing contracts.

- Given that structural separation has one clear advantage in a regulated context -- its ability to drive out anti-competitive conduct, there is no justification for prohibiting it on the basis of theoretical and unsupported conjectures about contracting failures. There is thus a presumption in favour of permitting voluntary structural separation and divestment.

- If problems of investment co-ordination were acute, separation would destroy shareholder value and would be irrational. In these circumstances, a regulator opposing separation (and ignoring its pro-competition benefits) would seem to be substituting its judgement over business policy for that of the firm's mangers and owners -- which is hardly justifiable.

- This leads to the conclusion that regulators should adopt a permissive approach to separation, even as they may have legitimate suspicions of the opposite policy of integration.

0 Introduction

It has long been recognised that vertical structures in industry can take many forms. The separate stages of production can be managed within an integrated structure, in which the co-ordination of the value chain is accomplished within a hierarchy of management. Alternatively, the structure can exhibit separation, and the relationships between the successive stages can be governed by contracts within a market setting.

Economic analysis first elucidated the advantages and disadvantages of the various alternatives within unregulated sectors, but the possibility of intervention by regulators in the structure of a regulated industry has attracted widespread and increasing attention. The focus here, in respect of both sector-specific regulation and competition law, has been on preventing the vertical transmission of market power via institutional bias and/or preferential treatment of affiliated upstream or downstream businesses. It is well established that under certain conditions, a monopolist in one part of the value chain, even when subject to price control there, has both the means and the motive to engage in discriminatory behaviour with respect both to price and non-price variables and with the aim of excluding competitors from downstream (or upstream) markets.

The classic behavioural response is to monitor and prohibit such behaviour, but this is often highly intrusive or ineffectual, giving support to structural remedies based on different forms of network separation, or network unbundling[1]. These remedies can range from accounting separation, through functional or operational separation[2] (creating separate business units, with an obligation to treat both affiliated and unaffiliated customers or suppliers in a transparent, equivalent manner), to full structural, or ownership, separation. In its 2001 Recommendation on structural separation in regulated industries, the Organisation for Economic Co-operation and Development (OECD) said that structural options should be considered on their merits. In other words, there should be no universal presumption in favour of an integrated *status quo*. It repeated this conclusion when revisiting the subject in 2006.

The fixed telecommunications sector is the subject of much discussion in relation to separation, largely as a result of the Office of Communications' (Ofcom's) acceptance of BT's undertaking in 2005 to implement the operational separation of its access network, or "LoopCo") business in the United Kingdom (UK), with the exception of Northern Ireland. The European Commission is also likely to propose the inclusion of operational separation as a remedy throughout the European Union (EU) as a whole from 2010, to be available in restricted circumstances.

These would be regulatory measures, imposed on operators. But what if an integrated operator wishes to divest a particular set of activities, which might be its access business (the remainder of its network and its retail business remaining together) or its retail business alone (leaving all its network activities in a single business)? The former proposal has recently been made by Telecom New Zealand; the latter is in contemplation by eircom in Ireland.

A standard objection to the imposition of structural separation in telecommunications is that it will, *inter alia*, impede the desirable co-ordination of retail, core network and access network investment activities. In fact, the same charge can be levied against functional separation

[1] Separation and unbundling (in their various forms) are often used as synonyms.; this paper uses the former term in relation to telecommunications.
[2] The former term- functional separation- is used throughout this report.

which, if it is to be implemented in a way which prevents discrimination, requires the 'ring-fencing' of both operational and investment activities undertaken by the separated entities, in order to exclude the possibility of abusive strategies. This is acknowledged in functional separation rules introduced or proposed by European regulators, such as Ofcom in the UK and the Communications Regulatory Authority (AGCOM[3]) in Italy.

This extends the scope of, rather than eliminates, the argument against separation based on problems in co-ordinating investment. How much validity does it have? This can be addressed from both a theoretical and a case study perspective.

The theoretical literature on separation and integration points to possible problems, which might arise in a separated structure if one party behaves opportunistically. A typical example of such behaviour might be for one party to wait until a partner has purchased a specialised asset, and then seek to lower the price paid for the output of such an asset, in the knowledge that its owner has no alternative outlet for its production. Vertical integration would solve this problem by aligning incentives in the upstream and downstream segments. But this examples hinges upon the parties' inability to devise a contract (for example, a long-term one) which would overcome the problems of such post-investment opportunism.

A further theoretical objection to separation is that if both stages of production exhibit market power, then the supplier of network services will take an excess profit, which will then be added to when the retailer imposes its excessive margin. But regulators seek to prevent precisely such exercises of market power by price control. Thus theoretical arguments against separation are based on pessimistic assumptions about both contracting and regulation.

To see whether these are justified, it is desirable to examine the balance of advantage between separation and integration in a number of illustrative case studies of both regulated[4] and unregulated sectors.

[3] Autorità per le Garanzie nelle Comunicazioni.

[4] In relation to regulated sectors, it must be acknowledged at the outset that creating the right incentives to invest has problems in many settings, which are not related to their structural features. As a simple illustration, rate of return regulation can lead to over-investment or gold-plating (also know as the Averch-Johnson effect), while the standard price-cap model (which a hybrid between incentive regulation, within a given price control period, and cost-based regulation at the renewal of the cap) encourages the under-fulfilment of exaggerated investment plans. It is vital to distinguish these problems of dysfunctional interaction between operator and regulator from any additional issues created by the choice of vertical structure.

1. Separation in industry

The idea of segmenting telecommunications networks or of separating them from retailing activities has been current since the mid 1980s, when in the United States the Bell System was broken up into a long distance and seven regional operating companies. Subsequent privatisations in Europe and elsewhere did not follow that pattern, but in recent years regulators and operators themselves have revived the notion of separation. The options under discussion now include:

- mandatory vs. voluntary separation;
- ownership (structural) vs. legal vs. functional (operational) separation;
- separation of the loop only (LoopCo), vs. separation of all network services (NetCo) vs. more complex (e.g., 3-way, or '3-box') separations.

The purpose of this paper is to focus on one major aspect of the separation debate – the linkage between separation and investment incentives. This is crucial because opponents of (some forms) of separation argue that difficulties in the co-ordination of investment across the separation boundaries thus created from a major or even an insuperable objection to its implementation.

Sections 1 and 2 set a context for what follows, in the form of a discussion of separation in regulated sectors and an application to telecommunications in particular. Section 3 reviews some economic models, which analyse the consequences of different forms of separation in more general contexts. Our discussion focuses on the subset of this work dealing with the effects of vertical structures on investment.

Section 4 examines how the issue of separation and investment has played out first in a number of mainly regulated sectors – railways, airports and energy, and also in non-regulated settings, including automobile manufacture, integrated circuits for personal computers and energy (again). These case studies show a range of more or less successful vertical relationships. Finally, Section 5 brings together the analysis of the previous sections to draw conclusions on the effects on investment of separation in telecommunications.

1.1. Separation and integration in sectors of the economy

Providing goods or services to end-users typically required a series of steps, including such things as procuring raw materials, installing capital goods, fabrication of parts, assembly, transport and retailing. How firms reach the 'make or buy' decisions which determine their structure and the degree of specialisation in the economy has been a staple of economic debate since Adam Smith.

The revival of this topic originates in a paper by Ronald Coase (Coase 1937)[5], where the discussion –as it was in the subsequent burst of interest in the 1970s and 1980s- revolved around unregulated sectors, or more accurately, sectors affected only by generic competition law, which in this period developed a keen interest in the leveraging of market power from one point in a vertically integrated structure to another.[6]

Privatisation in the 1980s and 1990s forced the integration/separation issues onto governments' agenda. When they came to sell their utilities they had a free hand over how

[5] Coase returned to the same issue in an article published 59 years later in 2006; see Sec. 4.6 below.
[6] In fact, both the break up of Bell in 1985 and the functional separation of BT in 2005 were accomplished under competition law.

to set up the deal – either in vertically integrated form, or vertically separated. If governments' advisers told them that the sale of an integrated firm would generate more revenues, they were entitled to wonder whether this arose from the greater efficiency of integrated structures or greater opportunities such structures offered for the exercise of market power. For example, in the UK, the initial privatisations (of BT and British Gas) were integrated; those for electricity supply industry and the railways were separated.

More recent debates have focussed more comprehensively on the motives for separation, and these are first summarised. Then an account is given of different forms of separation and finally of the views on separation of the OECD.

1.2. For what problems is separation a remedy?

If an investor-owned firm chooses to separate itself, it is likely that its management and shareholders believe that doing so increases its shareholders' net worth. Given that separation is unlikely by itself to enhance market power (as the reverse process of vertical integration might do), there must be a presumption in favour of allowing it to happen.
The issues associated with mandatory separation are quite different. If a regulator imposes it, the motive is most likely to revolve around eliminating discrimination and preventing foreclosure. Consider a vertically integrated incumbent providing a variety of narrowband and broadband services, untroubled by the presence of an alternative wireline-based access network, such as cable. Regulation is likely to be applied under the existing European regulatory regime in the form of mandatory access (at either cost-based or 'reasonable' prices) to some of the incumbent's assets, such as the local loop, wholesale broadband access, call origination, termination and transit, leased lines, and so on. Such 'pro-competitive' regulation at network level is seen as an increasingly viable alternative to 'consumer protection' regulation in the form of retail price controls.

The success of this approach hinges upon the appropriateness of the terms and conditions of access to the assets in question. If the incumbent can offer better terms to itself than to its competitors in downstream markets, it can exclude them from or weaken them in those markets.

Such discrimination can take two forms: price and non-price. Accounting separation is designed to ensure parity between transaction prices paid by competitors for access and accounting prices paid by the separated entity's downstream affiliate. Excessive prices will show up in excessive return earned by the access-providing business of the incumbent. A generalised margin squeeze might be illustrated by negative returns, or a deficit, in the downstream unit. Of course such 'parity' is not complete, in the sense that competitors' access payments are a genuine marginal cost for them, whereas the marginal cost of the same service to the vertically integrated incumbent is its marginal resource cost (the extra physical cost of producing one extra unit, translated into monetary units). The latter is likely to be much lower than the former, when access prices are based on long run average incremental cost, with mark up, while the production process exhibits economies of scale.

Accounting separation might deal adequately with price discrimination. But non-price discrimination, which might be the result of historic network design and to some degree unconscious, is a different matter. Much of the UK case in favour of the operational separation of BT rests on the proposition was practising non-price discrimination and was likely to persist in doing so.[7] The proposed remedy is a redesign of business processes -

[7] See Ofcom (2005a) and Cave et al (2006) and the sources cited there.

functional separation – to ensure precisely equal treatment – 'full equivalence '- for both internal and external purchasers of the same service.

Vertically integrated firms often, whether by accident or design, discriminate against competitors. In such circumstances, separation will confer a pro-competitive benefit.

1.3. Types of separation.

A major issue in establishing a separation regime is to specify the behaviour required. In one sense, the focus should be transactions on the boundary between the separated components, but the objective of achieving non-discrimination here may have to be supported by wider-ranging constraints on the separated entity.

Table 1 contains a specification of separation options varying from accounting separation underneath the 'ladder' to partial or full ownership separation at the top.

Table 1. Separation Options

6-Ownership Separation(in whole or part)
5-Legal separation (separate legal entities under the same ownership)
4-Functional separation with localised incentives and/or separate governance arrangements
3-Functional Separation (BS)
2-Virtual Separation
1-Creation of a wholesale division
Accounting separation

The focus here is on the six degrees lying above (and excluding) accounting separation (see Cave 2006). These options are now described in more detail.

Accounting separation itself entails separate profit and loss statements and balance sheets for the separate entities. This can be accompanied by the creation of a special wholesale (or otherwise named) unit, with a dedicated management (1 in Table 1). This will be responsible at a managerial level for the production and supply of the relevant products, but with no guarantee, at this degree of separation, of non-discrimination between affiliated and competitive access seekers. Such accounting separation has been a regulatory obligation on most EU telecommunications incumbents since 1998.

Under this regime, the regulator can make attempts to ensure some loose equivalence between services to affiliated units and to competitors. However, these efforts are hampered by two factors in particular:

- the absence of a clean target level of equivalence- an ambiguity which leads to opportunities for the incumbent to continue to discriminate:
- the fact that incumbent's network, IT system and business processes were broadly designed within the context of a fully integrated firm supplying end users directly, but not supplying access services to third parties; the historic situation was thus "discriminatory" at that time of market liberalisation, when access products were grafted onto the network through the adoption of special procedures and technological fixes; commercial motives then perpetuated discrimination, whether intentional or unintentional.

The next variant considered is virtual separation (2). This means the imposition by the regulator of an obligation to achieve full equivalence in the services offered to internal and

external customers without any physical separation of networks, signalling systems, business premises etc. Vertical separation thus requires, in effect, a reengineering only of the transactions boundary to achieve equivalence, but no change in the underlying production processes. Virtual separation might be achieved, inefficiently, by degrading the quality of services provided to internal customers (for example, the speed with which orders are transmitted and processed) or by upgrading services provided to external customers. Virtual separation is likely to be much less costly than more comprehensive 'physical' separation.

Virtual separation is the *modus operandi* of many European telecommunications incumbents at present, including eircom, given the obligations for non-discrimination imposed on them since in the 1998 'First Package' and the 2003 EU Regulatory Framework[8]. The key issue here is the actual and perceived feasibility of achieving full equivalence in such circumstances; both are important, since lack of trust in the arrangements will deter investments by competitors almost as severely as actual discrimination. Still, as this approach has not yet been tried in the context of achieving full equivalence, it is only possible to speculate about how it would work and what it would cost. However, its similarity to previous attempts to outlaw non-discrimination, which in several jurisdictions are regarded as failure, is a handicap.

The next step up (3) involves functional separation, which requires reworking of underlying business practices and not just changes at the transaction boundary, as with virtual separation. The aim is to segregate particular assets and other inputs within a separate unit, which then trades using identical processes with both internal and external customers in way that can be verified transparently.

However, the separation is not complete; otherwise, we would be observing something equivalent to full ownership separation. Instead, the firms' assets can be separated in different degrees, as noted in Section 2.2 below.

Reverting to the typology, the higher level of functional separation (4) involves incentives for senior managers in the separated entity, and/or separate governance arrangements. If externally imposed, this involves more detailed regulation not only of the transaction boundary and production processes but also of the relations of production of the separated services. The simple argument in favour of tailored incentives is that senior managers will otherwise maximise group shareholder value rather than divisional profits, as a means of personal advancement and a response to share options. This may lead them to practice discrimination against competitors whose success in downstream market would jeopardise group profit. To prevent this, managerial remuneration should be tied to divisional performance and (possibly) restrictions should be imposed on movement of senior staff from the separated unit to the group.

A further escalation of measures in a similar vein would require the creation of a divisional board with non-executive directors independent of the group, or of a special scrutiny regime to enforce separation. This could take the further form of legal separation (5), a regime in which a separate board is created and separate statutory accounts are filed -- all designed to emphasise and support the independence of the separated entity.

[8] As stated in the Access Directive (Article 10), non-discrimination requires that an undertaking found to have significant market power "applies equivalent conditions in equivalent circumstances to other undertakings providing equivalent services and information to others under the same conditions and of the same quality as it provides for its own services, or those of its subsidiaries or partners."

The final option (6) requires separate ownership of the separated assets. This could be incomplete, in the sense that the group might exercise partial ownership – but this option will be ignored. Instead, we will focus on wholly separated shareholders with no motive for discrimination.

Six separation options have been described with their associated internal behavioural rules. Running in parallel are enforcement mechanisms. These can be internal or external. For example, the integrated group can set up an independent complaints body, to investigate the conduct of the separated entity. Or the regulator can investigate and impose sanctions for breaches of license conditions or of undertakings. As in other areas of activities, an effective external enforcement system with a high level of deterrence can to some degree secure the achievement of goals which go against the grain of a company's or a manager's incentives. Equally, a well-designed incentive mechanism can relieve the pressure of enforcement.

1.4. The OECD Recommendation on structural separation in regulated industries

A clear indication of the potential importance of separation in regulated industries is provided by the adoption in 2001 by the OECD of a rare recommendation concerning structural separation in regulated industries.[9] It argued for a careful balancing of the costs and benefits of structural measures against the costs and benefits of behavioural measures.

In a report based on experience of the recommendation published in 2006[10], the OECD concluded that:

- the recommendation is still important and relevant, and should remain in place;
- the suggestion to balance benefits and costs still holds, as does the view that these will differ from sector to sector.

The OECD's summary assessment of pros and cons of structural policies is given in Table 2 below:

Table 2 The Pros and Cons of Different Separations

Policy	Advantages	Disadvantages
Ownership Separation	Eliminates incentives for discrimination; allows for lighter-handed regulation of downstream entities.	Potential loss of economies of scope; may require costly and arbitrary separation
Club Ownership (ie ownership by a consortium of firms)	Eliminates incentives for discrimination within club	Club may seek to exclude outsiders; may facilitate collusion; only effective in certain circumstances
Operational Separation	May facilitate control of discrimination and anti-competitive behaviour	Possible lack of profit motive reduces incentive to provide innovative and dynamic services

Source: OECD (2001)

The present report seeks to adopt the OECD's approach. It has an open mind towards separation; it examines comparators to the telecoms sector as well as that sector; it recognises the inevitable difficulty in speculating about alternative structures but seeks to avoid a presumption in favour of, or against, the *status quo*. The approach adopted here is

[9] OECD (2001)
[10] OECD (2006)

like that of the OECD: it is pro-separation only in the exceptional sense that it is willing to contemplate and evaluate non-standard structures.

The OECD recommends an appraisal of alternative structures on their merits. The present paper adopts this approach.

2. Separation in the telecommunications sector

This section gives an overview of recent debates on separations in telecommunications. It starts with a review of where to separate telecommunications network, continuing with a discussion of functional and structural separation in the sectors and then ends with a preliminary discussion of investment issues.

2.1. Where to separate

Separation involved drawing boundaries around particular activities and the assets, which they utilise; these decisions then generate separate accounts, separate 'divisions', separate legal entities, or separately owned firms.

In a complicated business like telecommunications there is enormous scope for choosing different points of separation, not least because the number of different activities involved is very large.[11] Under voluntary separation, the decision rests largely with the firm in question. Under mandatory separation, the regulator may have the last word, though some form of negotiation may be involved. In either case, the decisions are likely to reflect economic considerations relating to whether co-ordination issues should be tackled by markets under separation or within the hierarchy of an integrated firm, and also to the scope for leveraging market power from one point in the value chain to another.

When this latter consideration is deployed to justify some form of mandatory separation, the logic of the argument in Section 1.2 above suggests that separation should occur on the boundary between markets where the incumbent exercises persistent market power (and hence can discriminate with anti-competitive effect) and markets which are potentially competitive. It follows from this that the appropriate division depends upon current and predicted market developments. These will vary with the size of the economy which the telecommunications sector is serving: in a small country the scope for competition will probably be smaller.

There is also a crucial question of priorities. In some countries, proposals for functional separation may focus on current generation access networks - notably copper loops providing ADSL. In others, separation efforts aimed at achieving full equivalence between third party buyers of access services and the integrated firm's retail affiliate can be focussed on next generation networks (NGNs).

In discussions of separation, the two principal candidates for making a single split are to do so between retail and wholesale (hereafter called the "NetCo" model) and between access or

[11] Temin and Louis Galambos (1987) for the problems involved in separating the US Bell System in 1984.

the local loop and all non-access services including retail (hereafter called the "LoopCo" model)[12]. Underlying this is a three-way classification as shown in Table 3.

Table 3. Breaking up telecommunications services

Retail	Marketing and selling services to end-users and managing the end-user relationship
Network (non-access)	Core network services Call origination, termination, transit etc Trunk segments of leased lines Some backhaul
Network (access)	Unbundled local loops Wholesale line rental Some backhaul Tail segments of leased lines

Different (national) markets will exhibit different potentials for competition in these components. While retail is actually or potentially competitive everywhere, and the local loop is hard to replicate (except where ubiquitous cable television network(s) exists), the potential for competition in non-access network services is highly variable. In a large national market, such as Italy, France, Germany or the UK, there may not be persistent competition problems core network services. Other non-access network products may present harder, but not insoluble problems. Where a problem is likely to be found is in access services (in areas without cable or effective wireless networks). The intermediate function is backhaul, which will probably be (actually or potentially) competitive in some areas but not in others. In a small economy such as Ireland, the scope for developing competition may be much more limited, so a different 'optimal' point of separation is inevitable. It is worth pointing out that under structural separation the LoopCo variant imposes greater investment co-ordination problems than the NetCo model, as it places the access and core networks in separately owned companies. In the NetCo version, all network investment lies within the control of a single company.

The discussion so far has only included one separation into two 'boxes'. But more complex 3-box and above solution, are also possible. The New Zealand Government has proposed any access/network(non-access)/ retail split.

From the standpoint of achieving pro-competitive benefits of separation, the optimal point of separation depends on the scope for competition in the market in question.

2.2. Operational separation

We now examine a number of implementations of or proposals for operational separation.

A. The UK

The most talked about example of operational separation is that of BT – the assets separated, in a division known as Openreach, comprising BT's local loop. The separation emerged from undertakings offered by BT at the end of a 'market review' undertaken by the UK regulator, Ofcom, using its powers under UK competition legislation, the Enterprise Act 2003, rather than its sector-specific powers.

[12] Of course several separations can be made simultaneously, as in the New Zealand 3-box model noted below.

That legislation enables Ofcom to investigate a market, or set of communications markets, which appear to be subject to persistent competition problems. Thereafter Ofcom can take no action, accept undertakings from any firm or firms, or make a reference to the Competition Commission which, if it reaches an adverse finding, can impose remedies, including structural separation. Any breaches of undertakings given carry the possibility of a fine and can trigger legal action by injured parties.

In this case, BT undertakings were accepted by Ofcom and included functional separation of the access network. In more detail, the undertakings, expressed in a 55-page document (Ofcom 2005a), are, as follows, to:

- establish an operationally separated access services divisions (subsequently named Openreach), located on separate premises;
- ensure full equivalence for key access products by agreed dates;
- establish an Equality of Access Board (EAB) to police the undertakings;
- consult on the development of its next generation networks.

To date, an access services division has been established under the name of Openreach; fully equivalent services are available for a number of products; the EAB has been established; and collaboration on NGNs has progressed via an industry group called NGNUK. BT has also recorded its progress in meeting its key performance indicators.

The UK experience of functional separation is still in its infancy, but regulators claim to have detected a major change in the way BT's anti-discrimination obligation now operates, in particular, with increasing delivery run rates and quality improvements in the wholesale access products provisioned. Previously, it was unclear, responding to the injunction: "do not discriminate in a way which has a material effect on competition". Scores of complaints from competitors were investigated and inquiries by the regulator undertaken, not one of which led to a formal finding adverse to BT.

The new requirement in the undertakings, by contrast, is sharper – the 'bright line' of full equivalence of services supplied to internal and external customers – and for that reason much more easily verifiable and resolved in legal proceedings It is supported by a complaints body, which is still finding its way, and a senior management incentive scheme only recently completed. It is also combined with a major change in BT's strategy and rhetoric. The company now denies that it benefits in the UK from the advantages of incumbency.

Will Openreach stay within the BT Group, or will it be structurally separated in due course? At the analysts' conference following the publication of BT's results in May 2007, its chairman was asked this question, and he replied: 'we have not closed the door, and we will continue to examine all the options in relation to Openreach.'[13]

B. Telecom New Zealand

In May 2006 the New Zealand government undertook a stock take of the telecoms sector and its regulation, the conclusion of which was a proposal to adopt new rules on mandating access to Telecom New Zealand's (TCNZ's) assets and to impose a functional separation on the company. Following legislation in December 2006, the Ministry for Economic Development published a consultation document on the functional separation in April 2007[14].

[13] *BT Group plc- results presentation*, May 17 2007, www.callstreet.com
[14] MED (2007).

The document set out proposals for a 3-box model, in which TCNZ would be separated into separate access networks services (ANS), and wholesale and retail business units, with ANS operating on a stand-alone basis.

In the course of the consultation, TCNZ proposed a structural separation of its access business (discussed below). It made this offer in part as a result of its serious reservations concerning the government's separation proposal, which it regarded as being costly and over-intrusive.[15]

In May 2007, the Minister for Economic Development announced his decision to press ahead with the 3-box operational separation proposal, at the same time as he would give further consideration to the TCNZ proposal.

C. Other discussions of functional separation

Within Europe, at least two regulators, in Italy and Sweden, are giving thought to imposing operational separation on respectively, Telecom Italia and TeliaSonera. The Italian regulator (AGCOM) has conducted a consultation on the matter that ended in July 2007, and has announced the intention of imposing functional separation on telecom Italia by the end of 2007. The Swedish regulation (PTS) has proposed that legislation be passed by the Swedish Parliament which would implement separation speedily.

At EU level, there has been a long debate over whether functional separation should be included as an additional remedy in the new access directive, expected to be legislated in 2008/9 for implementation in 2010. In June 2007, Commissioner Reding hinted strongly that she would propose such a measure.

The UK functional separation has been accomplished by BT in a whole-hearted and constructive fashion, as its competitive rivals acknowledge (Ofcom 2006). But if a company on which functional separation was being imposed were to resist the process, it and the regulator would be locked in a long-term battle over whether the company was abiding by the rules. A battle over structural separation is, by contrast, necessarily time-limited, as it ends with the separation.

A regulator-imposed form of functional separation already operates – with uncertain benefits – in the UK, and other regulators are beginning to follow suit.

2.3. Structural separation

There are many examples of relatively uncontroversial divestments by telecommunications companies. Some have been voluntary- for example the sale of mobile operations by eircom and BT; some have been made or offered as part of undertakings associated with proposed mergers- for example the undertaking offered by a proposed acquirer of Portugal Telecom to sell either its PSTN or its cable network; some have in effect been imposed- for example the sale by Deutsche Telekom of its cable networks. Generally, these divestments have the characteristic of being horizontal in nature, involving the separation of activities supplying sometimes competing services to end users: local and long distance telephone calls; fixed and mobile services; services provided by telecommunications and by cable television networks.

[15] TCNZ (2007a and b).

By contrast, there is no working example in Europe of full structural separation of the kind described in Section 1.3 above – either the LoopCo or the NetCo model. We are therefore forced back on discussion of the recent Telecom New Zealand proposal and on *a priori* analysis, such as that published by the OECD in 2003.

In April 2007, TCNZ proposed the creation of a structurally separated company which would own TCNZ fixed line local bottleneck assets (TCNZ 2007).[16] Telecom proposed an explicit regulatory contract between the New Zealand government (not the regulator) and itself covering what services would be provided, what prices would be charged for them and what investments would be made.

The remaining (non-access) component of TCNZ would not be operationally separated, in the manner proposed by the government in its 3-box model, into wholesale and retail divisions. The logic of this is presumably, that with a structurally separated access operation, the residual risk of market power being leveraged from the core network to retail is insufficient to justify the costs of operational separation. TCNZ proposed that equivalence rules would apply to specified wholesale services.

The identification of access assets contained some disputable elements. Notably, TCNZ proposed that the DSLAMs in the street cabinets (which fibre would reach in TCNZ's proposed next generation access network) would remain with TCNZ's wholesale division, rather than pass to the separated access network operator. Competitors would then be able to lease them from TCNZ, via the access company. This reflected a view that it would be uneconomic for competitors' to install their own DSLAMs in street cabinets.

This quasi-voluntary proposal by TCNZ follows a more general review of the costs and benefits of structural separation of the local loop published by the OECD in 2003. These were summarised in Table 4.

Table 4. Some benefits and costs of structural separation

Benefits	Costs
Benefits of altered incentives towards non-discrimination	High implementation costs
Improves information and restricts cross-subsidisation if LoopCo is restricted from providing services	Irreversible
Reduces anti-competitive activities, leading to increased competition and its benefits	Negative impact on broadband development
Relatively simple (compared to behavioural remedies)	Loss of economies of scale and scope
Management and regulators can focus on the wholesale network	Erosion of incentives to upgrade network
Reduces the need for and cost of regulation	Loss of bundling advantages

Source: OECD(2003), p.27

[16] Confusingly, this is referred to as a NetCo proposal, whereas in our terminology it would be seen as a LoopCo option.

The OECD further noted that :

> "It soon becomes clear that even a limited quantitative assessment of structural separation (that ignores secondary benefits and costs) will be difficult. Attempting to identify the benefits and costs of structural separation involves assigning values to many current benefits and costs that are difficult to quantify, as well as projecting them forward into a future in which telecommunications technologies and markets may change in unpredictable ways.
>
> Clearly, an *a priori* assessment of the benefits of structural separation will depend unavoidably upon judgements/assumptions made on a range of variables, including the competitive constraints now placed on a vertically integrated incumbent (by regulation, by the competing infrastructure and also by other infrastructures deployed in the future)."

The OECD concludes that there is little evidence that the benefit of mandatory structural separation of the local loop are sufficiently in excess of their outputs to justify the risk,

The paper is, however, relatively silent on the costs and benefits of a network/retail split; it has little of substance to say on voluntary separation (there being only one example from the United States at that date); and it does not incorporate in its baseline case the possibilities of mandatory functional separation. The last point is significant, in that if regulators are going in any case to enforce functional separation, the incremental costs and benefits of structural separation are quite different than if the alternative is full integration.

The OECD's 2003 objections to mandatory structural separation do not apply to voluntary separation and fail to take account of the emergence of functional separation.

2.4. Investment under different types of separation

In a section discussing what incentives should be put in place to ensure the necessary investment and innovation to upgrade the local loop: the OECD paper on structural separation notes the following (OECD 2003, 24):

> "There are concerns over whether there will be adequate investment in network infrastructure when providers are denied the revenues and consequent incentives that flow from vertical integration. This problem is acute in the telecommunications industry, where technological change is rapid and where investment demands are pressing. Problems of co-ordinating investment between the wholesale and retail operators would also impede investment and innovation. These problems could be considerable and could serve to delay the extension of fibre closer to the customer.
>
> It is not evident that ownership separation would result in greater innovation. Structural separation is likely to have some success in regard to promoting competition and this in turn could promote network enhancement. But some analysts ... consider that in many OECD countries, the changed stock market sentiment towards telecommunications operators may well mean that it is the incumbents with the benefit
> of a steady cash flow that could be in the best financial position to enhance the network and the local loop. And structural separation threatens to weaken the ability of incumbents to make such crucial investments."

Clearly this last point would not apply to the NetCo model, which deliberately creates a company with a steady cash flow for the purposes of raising capital at favourable rates.

Later the paper notes (p. 29) :

> "A vertically integrated telecommunications company may achieve lower cost structures, for instance, by spreading billing costs across a wide range of services. Similarly, it can produce service packages ("bundling") at a lower cost than a firm producing the same services on a stand-alone basis. Vertical integration enables the firm to co-ordinate production and investment decisions by minimising external transaction processes and their attendant costs and delays. Such a mode of operation is particularly necessary in an industry operating on the 'technological frontier', where internal processes and structures need to be highly responsive to change.
>
> Many of these potential sources of cost efficiencies can be at least partially exploited through contractual arrangements between separate firms. Thus, an understanding of the costs of separation requires a comparison between the cost efficiencies achievable under integration and the cost efficiencies achievable through contractual arrangements. To the extent that there are vertical contractual arrangements that can achieve the same efficiency benefits as integration, the economies of scope are accordingly diminished. And, of course, any diseconomies of scale and scope should also be recognised."

General issues of co-ordinating investments across a transaction boundary and the adequacy of the contracting alternatives in comparison with a vertical integration are considered in the next two sections. But it is useful to discuss here the difference in the investment decision-taking regime between operational and structural separation.

Consider first the UK arrangements for Openreach. These are set out in the lengthy UK undertakings, which give Openreach considerable latitude, recognising that incentives are not enough: management requires some latitude to pursue its objectives. But equally, it must not be too much or shareholder control mechanisms are endangered.

This conflict comes to a head with investment decisions. The well-known precedent in Australia of Telstra's Foxtel network is an illustration of a predatory access investment. As the chairman of Telstra then acknowledged, the investment was expected (or intended) to lose money in the supply of broadcasting services, but to be profitable overall by virtue of defending the company's telephone revenues from a competitor which sought to provide both telephone and broadcasting services on a single network. Other circumstances might turn a separately profitable access project into deficit if repercussions in other markets were taken into account.

The BT Undertakings seek to deal with this point by imposing the following condition (Ofcom 2005, para 5.13.2).:

> "Any investment decisions required in consequence of the product road maps and volume forecast referred to in Section 5.13.1 shall be considered solely on their own merits, and shall not take into consideration the potential impact on other product offered by BT's businesses downstream of AS [Access Services- ie Openreach] other than in as much as they affect aggregate demand forecasts."

This paragraph crystallises an acute problem. On one hand separation is intended to allow the separated unit to make independent decisions; on the other hand, it is felt that major strategic decisions should be taken at group level. Yet the Group's incentives are by definition group-wide – and may thus threaten the purity of the motives encouraged by 'neutral' local incentives arrangements.

Under the UK arrangements, the Equality of Access Board (which 'polices' BT's Undertakings):

> "shall also be responsible for monitoring and reviewing the product roadmaps and volume forecasts – as well as the associated investment decisions as referred to in Section 5.13.2 [quoted above], as they relate to AS and SMP (significant market power) products."

According to a BT official, the issue has not yet been faced, in large part because BT has yet to develop plans for a next generation access network, which would represent a major strategic investment in access. But the dilemma is an acute one: either the functionally separated entity acts as if it were structurally separated, or it remits strategic decisions to a group level at which interrelationships between the component parts are taken into account. As a senior Cable and Wireless executive recently put it:

> ''The creation of Openreach is fundamentally flawed because if Openreach is to work properly, the way the regulator wants it to, you are asking the main board directors of BT to make decisions that are not in the best interests of shareholders. I find it odd that anyone would have thought it could possibly work given that structure and the fundamental conflict involved."[17]

The situation, which would be created if the functional separation proposal consulted on by AGCOM were implemented, is even clearer. That proposal requires a completely independent investment strategy by the separated LoopCo, carried out by a board with a majority of members appointed by the regulator.

As far as investment co-ordination is concerned, the difference between operational and structural separation diminishes on inspection; it is likely that objections to the one form of separation will apply to the other.

2.5. An interim summary

In the first two sections this paper has described summarised discussion and experience of separation within (mostly) regulated sectors.

The key points to emerge are:

- **separation has a clear benefit – the elimination of the potential for anti-competitive behaviour**
- **it may also have costs (loss of economies of scope, co-ordination costs, etc.)**
- **the benefit and costs are likely to vary from sector to sector**
- **operational separation is a comparative novelty in the telecommunications sector, and experience of structural separation is even more limited;**
- **in relation to the co-ordination of investment, the two types of separation appear in practice to impose similar, if not identical challenges.**

[17] *The Guardian*, 29 May 2007, p. 26

3. Economic theory of separation and integration

The structure of the telecommunications sector arose in the context of public ownership in Europe and high levels of regulation in North America, and, like many utilities, it took the form of a monopoly vertically integrated between its upstream and downstream elements. The upstream elements comprise network infrastructure, which can be broken into two large sub-elements: the core network and the (local) access network. The downstream element comprises retailing operations, covering the customer management relationship – chiefly marketing and billing. In this section, we address the theoretical arguments for and against separation in its various guises.

3.1. Economics of vertical integration

Vertical integration or the combining of upstream and downstream elements is at the heart of what Coase (1937) termed *industrial organisation*. Coase's original paper on the nature of the firm inquired into what determines the boundary between the firm which organises production internally, and the market where the price mechanism operates through contractual exchanges. Coase suggested that at some point reliance on the market for exchange would involve costs of negotiation (what we now term transactions costs) that would exceed the costs of organising the same outcome internally through direction and authority within a firm. Coase also noted that the boundaries between the firm and the market could change, as markets and technologies change.

Later writers developed Coase's ideas on the boundaries of firm and articulated in greater detail a property rights approach founded on contracts.[18] Klein, Crawford and Alchian (1978) -hereafter "KCA" - extended the debate to the specific issue of vertical integration, which is of primary concern to our investigation.

KCA explicitly deal with the problem of opportunistic behaviour (also called 'hold-ups'), which were recognised as an important component affecting the organisation of economic activity. They show that as assets become more specific, the scope for opportunistic behaviour grew.

Assets might be specific for a number of reasons.[19] They might be site-specific, as a local loop is located to serve a particular area. Their specificity might be due to their functionality- for example they might only be useful for highly specialised purposes, and for that reason lack resale value. Or they might be dedicated to producing goods or services for a particular buyer. In some cases one contracting party has specific assets. In others they both do- in other words the assets are co-specialised.[20]

Where assets are specific, in order to recover their costs, firms have to earn more on the assets in their intended specialised use than they would be able to earn in another use. This surplus, known technically as a quasi-rent, is, however, subject to appropriation or hold-up by their contracting partner. Because one party (at least) will fear appropriation if it invests, the investment will not occur in the first place. For this reason, separation is said to lead to co-ordination failures in investment. KCA claim as an empirical regularity that the lower the appropriable quasi rents, the more likely transactions will rely on a contractual relationship rather than common ownership.

---

[18] See Hart (1995) and Milgrom and Roberts (1992) for an excellent summary of the literature on these topics.
[19] These distinctions are set out in Williamson (1985) pp 95-6.
[20] See the example in Sec 4.8 below of a mine and a mine-mouth power station.

The relevance of this to telecommunications can be shown by way of example. Assume an asset is owned by U (this could be the upstream element such as the local loop) and leased to D (the downstream element such as the retailing operation). The quasi-rent value of the asset is the excess of its value over its salvage value – where the salvage value is the value in its next best use to another renter.

Suppose D leases access to the loop at €10,000 per day and the amortised value of the loop is €6,000 per day. The loop cannot be moved from its location and hence has no salvageable value elsewhere. Operating costs are assumed to be €4,000 per day and paid by the loop owner who supplies the wholesale services to the retail company. Suppose another retailing operation E is willing to pay a most €8,500 per day. The quasi-rent on the loop is therefore (€10,000-€4,000) = €6,000, that is the revenue minus operating costs minus salvage value (which is zero). The quasi-rent from retail company D relative to retail company E is €1,500 (=€10,000-€8,500).

At €10,000 revenue per day from the retail company D, the loop company would break-even and obtain a fair return – which in practice would be the likely result of price regulation. But once the loop is installed the retail company could offer to pay only €8,500 and the loop company could do no better by offering the service to retail company E. The effect of this would be the retail company would appropriate €1,500 of the quasi-rent from the loop company. Fear of this eventuality will prevent investment in the local loop.

The above numerical example is adapted from KCA. They assert that vertical integration is a means of economising on the costs of avoiding risks of appropriation of quasi-rents. Another resolution to the problem of opportunism of this kind lies in the shape of contracts; KCA dispose of this option by supposing that such contracts would be too costly to write (due to specifying all important elements such as quality) or are unenforceable. The following section provides important counter-examples.

If contracts are unenforceable or cannot be written due to the difficulties of specifying all contingencies, another solution is regulation.[21] Unusually this may take the form of a price-floor and a price-ceiling. On the one hand, regulation is required to counter market power held by the loop company, requiring a price-cap to prevent monopoly abuse. On the other hand, a price-floor would be required to prevent *ex post* opportunistic behaviour by a retailing company (as *ex post* the retailer would possess some degree of market power).

A contractual solution needs to safeguard against conduct by a LoopCo (such as degrading quality, which lowers operating costs) and against opportunism by the retailer or service provider. If the latter are especially differentiated (notably in terms of value between the highest and the next highest company), and their fixed costs are substantial, then either a sophisticated contract is required which protects them from appropriation , or common ownership may emerge. But if retailing is competitive, then common ownership will only be partial. And retail competitors of the integrated firm will be doubly reluctant to invest, since their rival will probably have an additional motive for abusing them, i.e., the prospect of excluding them from the market.

Another transaction cost stems from *ex ante* problems of information misrepresentation[22]. The problem of information misrepresentation is another factor that may motivate vertical integration rather than separation at the boundary. But another factor, which may be

[21] See Crocker and Masten (1996).
[22] Barzel (1982).

responsible for vertical integration, is market power. If market power is exercised both upstream and downstream, this can give rise to the problem of double marginalisation.

In some sectors, market power often resides in more than one part of the value chain and this could in theory give rise to a problem of double marginalisation. This might arise if the network company and the retailing company each possessed market power and each separately added an excessive profit margin. To resolve the problem of double marginalisation (which is a form of externality effect), vertical integration or common ownership is a solution. However, franchising is also a solution – for example a two-part tariff where the unit price is equal to marginal cost and the franchise fee is monopoly profits would restore the monopoly price. If the network company sets the two-part tariff within a regulatory setting, then rather than monopoly profits, normal profits may be obtained.

However, the fear of double marginalisation is misplaced because regulation will be employed to prevent the emergence of excessive margins. The notion that monopoly power could rest unchecked in telecommunications is unrealistic and so it is unclear how relevant the double-marginalisation concept is in regard to the problem we are addressing. In practice a combination of regulation to check monopoly abuse at network level , and (possibly small numbers) competition elsewhere is much more likely.

Adopting another approach in the theoretical literature, Cyrenne (1994) has shown, in a game theoretic model using the perfect equilibrium concept that when products in the downstream market are close substitutes (such as communications products), manufacturers will find it preferable to vertically separate. The intuition is straight forward: as products become more alike upstream manufacturers wish to exploit the flexibility of being able to sell to whomever they choose in the downstream market. In effect the externality impact giving rise to double marginalisation is eradicated as the market downstream becomes more competitive.

Cremer, Cremer and De Donder (2006) - hereafter “CCD” - have extended the debate recently to consider the impact of vertical separation (what they call “unbundling”) on the incentives to invest in network size. As discussed above, there are a variety of forms of separation ranging from the weakest, accounting separation through to ownership separation. CCD consider two possibilities – legal separation and ownership separation. Under legal separation the upstream firm will determine its network investments taking account the reaction of the downstream firms, even though it may not control the actions of the legally separate entity downstream. Under ownership separation, the network is only interested in achieving the best outcome on the assets it directly employs.

CCD investigated legal unbundling by looking at the upstream firm, which first chooses the size of the network, and then at two downstream retailers who compete using the network. The upstream firm is not allowed to discriminate in legal unbundling but chooses the size of its network to maximize the joint profits of the network and the retail operations it owns. They show that vertical separation in the form of legal unbundling (where the upstream firm does not control the pricing decisions of the downstream firms) reduces network investment below the social optimum and correspondingly increases the retail prices downstream. The upstream firm anticipates the decisions of the downstream firms and reduces the network size to force up downstream prices (by limiting capacity) and increase profits.

Under ownership unbundling the upstream firm chooses network size to maximize its profits only. In this scenario, it achieves profit maximization by lowering the size of the network further than in the case of legal unbundling. This is because it has to deal with the problem of double marginalisation – whereby the downstream entities apply mark-ups over the

transaction price it determines. However, this outcome (and the double marginalisation- rests on the lack of any price control upstream. If network prices are adequately regulated, the problem with separation can be resolved.

The theoretical arguments against ownership separation assume that contracts cannot be written which will deter opportunistic behaviour once investments have been made, and that regulation will be unable to prevent the exercise of market power, so that separation will lead to double marginalisation. Each of these assumptions is highly contentious and deserves testing in practice; this is undertaken via the following illustrative case studies.

4. *Investment incentives with integration and separation: case studies*

A. Regulated sectors
This section examines how other regulated sectors than telecommunications have fared with investment, in both separated and integrated structures. The sectors considered are: railways, airports and energy. This is preceded by a short account of the impact of regulation on investment.

4.1. Investment incentives under price regulation

Where price regulation is implemented to prevent abuses of market power by a monopoly supplier, two broad 'ideal types' are available – cost plus and incentive regulation[23].

Under cost plus or rate of return regulation, allowable invested capital earns a specified rate of return, and that return, together with depreciation and operating cost, is fed into prices. The firm's incentives thus depend crucially upon the relationship between the allowable rate of return and the firm's cost of capital. If the allowable return exceeds the cost of capital, the firm will seek to invest without limit; this 'explains' gold plating in regulated utilities and is also known as the Averch-Johnson effect[24]. If the allowable return falls short of the cost of capital, investment ceases.

At the opposite extreme, under very high powered incentive regulation, the regulator would set an exogenous trajectory for revenue or price, and the regulated firm's incentive would simply be to minimise costs. This would induce efficient investment decisions.

However, such an indefinitely long price or revenue cap might lead to unsustainably high (or low) profits. Accordingly, regulators have concentrated instead upon a hybrid system which involves three to five year periods of incentive regulation (a price cap), punctuated by periodic re-basing of prices on costs, as the cap is revised. When a cap is reset, the new initial price can either be equated with current costs, or a glide-path can be constructed under which regulated prices converge on projected costs at the end of the period.

The mechanics of this process, in standard form, are as follows: the firm proposes an investment level for the impending price cap period which (it says) will allow it to meet demand at regulated prices. After scrutiny of such proposals (either individually or in aggregate), the regulator adopts a price cap for the period, which typically permits recovery

[23] For a recent survey of the general issue, see Guthrie (2006).
[24] Averch and Johnson, (1962).

of the annualised costs of new investment, as well as of existing assets (the regulated asset base) and of operating cost.

Suppose the regulated firm is able to meet demand without undertaking the full investment projected. Then for the period of the cap it saves the cost of capital and depreciation of investments not undertaken, thus augmenting its profits. However, the benefit survives only until the end of the price cap. At that moment, its regulated asset base is less than had been projected and it is the lower or actual figure which goes into the computation of cost-based price for the next period. Nonetheless, the benefit may be considerable. For example, in the England and Wales water and sewerage industry, the real allowable rate of return is about 5%, and assets have an average life of, say, twenty years. Suppose 'saved' investments are spread equally over a five year price control. Then the average annual profits accruing to water firms from this process of underinvestment are:

[5% (allowable returns) + 5% (allowable depreciation)] x [2.5 years x £4bn (size of annual investment programme)] = £1 bn.

Vigorous efforts have been made to avoid this consequence, without falling into the opposite trap of encouraging the unnecessary investment wastefully to go ahead. An example is for the regulator to offer firms a menu of increasing investment levels, associated with progressively lower allowable rates of return. By this means the regulator in effect buys out the firm's private knowledge. [25]

For our purposes the important point is that this underlying problem is quite independent of the vertical structure of the industry. In the UK, it is observable in relation to price-capped activities in the vertically integrated water and sewerage industry and in the vertically separated railway and energy sectors. It is thus a consequence of rational behaviour of any regulated firm and unaffected by structure. This means that there is a significant problem of distinguishing behavioural and structural effects, to which we return below.

Creating incentives for efficient investment is a major problem under existing systems of sectoral regulation. These well-known difficulties are independent of the vertical structure of the industry and should not be attributed to it.

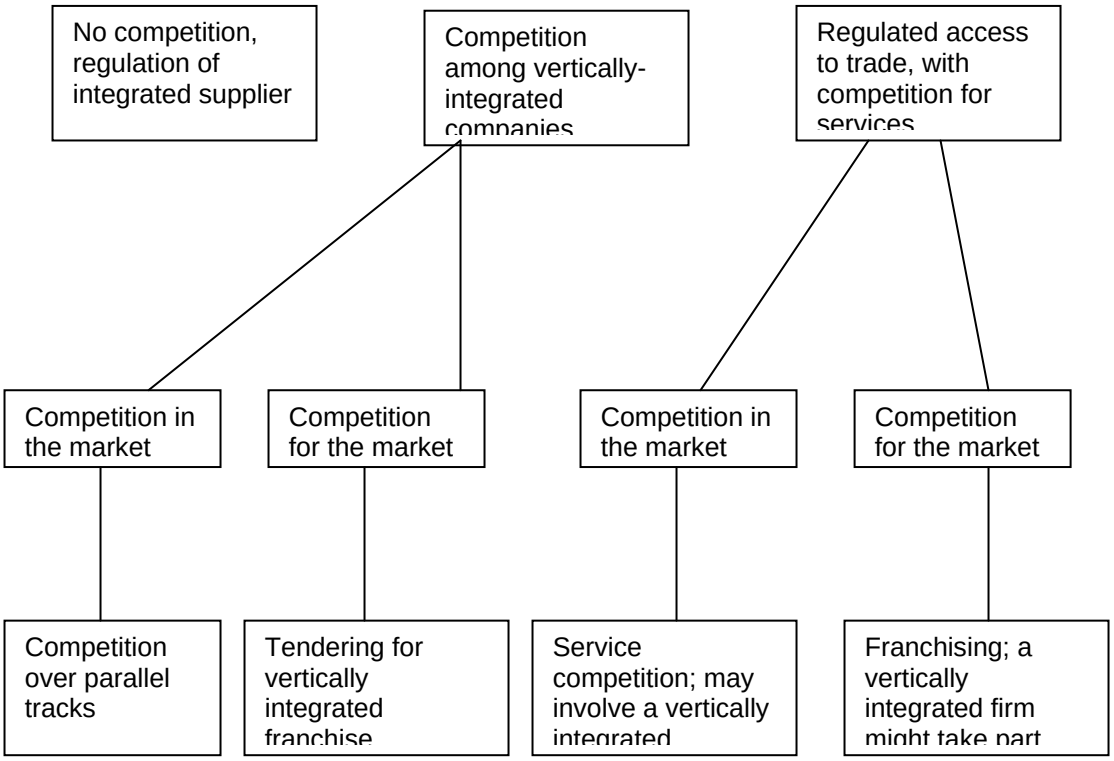
4.2. Railway sector

The railway industry, especially in the UK, is often regarded as the test-bed on which structural separation has failed. It is, however, arguable that special factors were at work, which made the situation unusually bad. Our chief interest is how investment incentives were affected.

The options for railways competition are set out in Figure 1, adapted from the OECD (2005).

[25] See Laffont and Tirole (1993). The Office of Gas and Electricity Markets (Ofgem) in the UK has used this approach; see Crouch (2006).

Figure 1. Forms of competition in railways



After considering various options, including vertically integrated regional companies and a hybrid structure, with some services integrated with the track, the UK Government in 1996 adopted a four-way separation of the passenger railway industry into: a regulated track and stations company; a large number of franchised train operating companies (a small minority of which compete over the same or 'parallel' tracks); putatively competitive rolling stock leasing companies (ROSCOs)[26]; and track maintenance companies. The freight industry operated on a separated and competitive (but highly concentrated) basis.

The potential operational problems in this structure were considerable. For example, maintenance companies, bore the burden of repairs to the track, according to their contracts, but could pass the cost of replacement through to the track company. Another example: it is well known that 'the point where steel wheel meets steel rail is about the size of a dime, but bad profiles as one or both can lead to millions of dollars worth of problems for railroad car and maintenance of we people'. It is asserted that national track infrastructure was damaged by train wheels with flat spots or them and that flat spots were caused by poor train maintenance by the ROSCOs[27], which were not liable for damage to the track.

The investment record of the UK track company, Railtrack, has also been subject to examination.[28]In the period after privatisation in 1996, investment fell short of expectations as the value of Railtrack's stock rose- largely as a result of too generous access prices set by the government. The regulator, the Office of Rail Regulation (ORR), initially left issues of

[26] In 2007, the UK Competition Commission is conducting a 'market review' of the ROSCOs, following complaints that they overcharge.
[27] Pittman, (2004).
[28] Gomez-Ibanez (2003) Ch 11.

capacity expansion to negotiation among the parties, but these proved difficult given the short duration of the franchises awarded to train operating companies and the different benefits to operators running trains on the same track. As a result of these difficulties, investment decisions were taken out of the hands of Railtrack and its successor, as illustrated by the Department of Transport's announcement in July 2007 of a 30-year plan for the railways, the costs of which will be largely defrayed or guaranteed as well as determined by the government.[29] Gomez-Ibanez (1993, p297) concludes that in the light of the virtual abandonment of plans to encourage competition among operators- which were thought likely to increase the subsidy budget, it would have been better , in the light of the co-ordination difficulties, to have maintained, or to revert to, an integrated structure.

However, this is predicated on the abandonment of competition at the service level. If this were not done, then the vertically integrated firm would have had to co-ordinate with its competitors, which would presented both co-ordination and competition problems. To put it the issue in a telecommunications context, if all prospects for retail competition were abandoned, then the argument for either voluntary or mandatory separation of network and retail would be much weakened.

This UK case study shows the possible drawbacks of imposing separation in the railway sector. But how representative is it, even of that sector? One approach is to examine the evidence of a cross-section of countries using econometric methods. One such study finds that organisational (something akin to functional) separation produces the same effects as full integration, but that structural separation has a positive effect on efficiency in a sample which, for data reasons, excludes the UK.[30]

The OECD has also produced a careful analysis of the impact of structure or investment. It reports that upgrades to the rail infrastructure in a vertically separated structure will confer benefits and impose costs on different operations to different degrees.[31] They may therefore require either protracted negotiation or a decision (as in the case of airports considered below) by a regulatory authority. The OECD notes that such co-ordination problems are present even in a partially vertically integrated structure, where the infrastructure provider is responsible for some services. They are wholly avoided only in a vertically integrated monopoly.

There is thus no consensus view on separation of the rail industry. According to the OECD (2005), where access is mandated, "decisions not to separate [structurally] should only be made after careful consideration of the costs that will result in the form of the additional regulatory burden and on-going residual discrimination." It goes on:

> "Beyond the principles identified above it is difficult to derive further clear principles that might guide decisions over structural arrangements in rail. For example, in the case of a service which is both a "dominant" user of the track infrastructure and where there is scope for effective competition in the vertically integrated entity or through vertical separation, and regulation of access to the track infrastructure. This might be the case for, say, one of the coal-dominated rail lines in Australia. These lines are profitable and could probably sustain effective competition, especially if train operations were separated from the track infrastructure. On the other hand, such separation could risk creating problems in regulating quality and investment in the track infrastructure and could undermine the ability of these services to form part of a

[29] Department of Transport (2007).
[30] Friebel, Ivaldi and Vibes (2004).
[31] OECD (2005).

seamless just-in-time “production line” coal-mining process. The appropriate choice of policy is not clear.”

Judgements about vertical separation in railways are coloured by the performance of the flawed structure adopted in the UK. Across a broader sample of countries, opinion is more divided.

4.3. Airports and airlines

Airports and airlines are an interesting example of the problems of co-ordination across an ownership boundary. This arises because vertical integration is not currently a normal option in Europe. Although historically both airports and one national airline – the flag carrier – have been government enterprises in common ownership, they were under separate management. Despite this, charges of discrimination were frequently levied by other airlines.

There are also issues concerning the horizontal concentration of airports. In Ireland, the government raised the issue of allocating an additional terminal at Dublin Airport to another operator than Dublin Airports Authority. Later this option was abandoned, despite loud protests from Ryanair, a major user. In the United Kingdom, the Competition Commission is undertaking a market review under the Enterprise Act of airports in the South East of England and in Scotland, which might lead to divestment in the South East of one or more of Heathrow, Gatwick and Stansted. [32] Simultaneously the Competition Commission is reviewing proposals from the Civil Aviation Authority (CAA), the airport regulator, for a price control for those three airports (and for Manchester Airport) for 2009- 2014).

Another unusual feature of the airports case is that significant investments – another terminal, another runway – tend to large and lumpy, and to promote considerable hostility from local residents which both delays and complicates the process. This places a premium on regulatory interventions to speed up investment. Moreover, airlines owning existing slots at airports where they are scarce may have a preference for postponing investment. This is because with landing charges regulated, they are the likely beneficiaries, through higher ticket prices, of excess demand for flights.

We may thus find that airports, driven by incentives of the type described in Section 4.1 above, have an incentive to gold plate their investments, and to bring them forward, while airlines, which have a concentrated structure, seek both to resist gold plating and, in certain circumstances, to restrict expansion.

These factors are illustrated by current UK policy. The government seeks sustainable development at Heathrow, including construction of a third runway.[33] The Office of Fair Trading(2006, pp5-7) – the competition authority – asserts that the British Airport Authority (BAA) has an incentive to make investments justifying higher charges to airlines without necessarily expanding capacity (‘gold plating’)’.

The CAA (2006) sought to place the burden on producing an agreed investment plan for the airports on the airport operator and the airlines themselves. To that end, the CAA has orchestrated a process of constructive engagement between the parties, which has produced results in some but not all contexts[34].

[32] This factor also makes horizontal separation more attractive, if it is likely to precipitate an ‘investment race’- see Guthrie (2006).
[33] Department of Transport (2006).
[34] CAA (2007) paras9.13-21.

It has also mooted the idea of linking revenue within the period of the price cap to investment.[35] It is, however, largely silent on how the proposed tie would work. One possibility would be to require a particular level of investment to be undertaken, or a threshold to be passed, before a price increase can be implemented. Alternatively, the precondition might be an increase in rated capacity. In either case, some kind of monitoring would be required- in addition to what has been needed in the past.

The discussion in Section 4.1 above has indicated some of the problems. Suppose first that the airport regulator and operator share the same information when the cap is being set. Then monitoring by the regulator will prevent gold plating. But there is still the need for an incentive for the operator to strive to economise on an investment project as it is being implemented. The current procedure (of allowing *planned* investment to be remunerated in the current price control period) provides such an incentive. It need not on balance reward the operator, since in-period 'news' can be bad as well as good.

Now suppose that information at the start of the period is asymmetric, so that the operator can exaggerate its requirements for investment. The situation here is more complicated, as on one hand, the regulator does not want the operator to be rewarded for exaggeration, but on the other it does not want it actually to incur wasteful expenditure, or any other expenditure which is unnecessary. In these circumstances too, it may be more efficient to discourage actual expenditure even at the expense of rewarding exaggeration. In other words, this may be a case, quite common in incentive regulation, where the regulator has to 'buy out' the regulated firm's private knowledge, while striving to minimise its scope by other means, such as improving its own data and understanding.

In the UK, the CAA may have a compromise procedure in mind, involving investment milestones, which overcomes these problems. According to the Competition Commission, the BAA's current price cap involves a small number of triggers, which reduce landing fees if milestones are not met and, even more, are contemplated for the price review currently under discussion from 2008 to 2013[36]. Amending the price cap to deal with a step change in investment might be both feasible and desirable, if it did not distort BAA's choice.

Thus a number of lessons arise from this discussion – including the possibility of seeking agreement on investment between the airport operator and the (downstream) airlines, and of a tie between in-period revenues and investments. But the problems of investment have not proven to be insoluble under separation, and there is no suggestion at present that public policy objectives could be furthered by integrating the airport operator and the largest airline and requiring that firm to offer airport access to competitors.

Problems characteristic of all regulated sectors apply to finding optimal investment incentives for airports, which reflect customers' needs for expansion. These have led to a number of palliative remedies, including constructive engagement and investment triggers.

4.4. Regulated energy markets

Vertical separation in the regulated energy sector provides a range of complex options, which differ in significant ways from electricity to gas. Ignoring the latter difference, the elements are: extraction/generation, transmission/distribution and supply/retail.

[35] CAA (2006) paras 6.18-6.20; CAA (2007) paras 9.74-5..
[36] Competition Commission (2007) pp5-6.

In the United Kingdom, the changes in industry structure over the past 20 years have been enormous. At privatisation in the late 1980s, a number of local distribution and supply (i.e., retail) companies co-owned the high voltage transmission system, which received energy from a highly concentrated generation sector. Since then, the transmission system has been separated from local distribution (and, more recently, has acquired and then sold part of the gas distribution network); generation has become much less concentrated, and generation capacity is owned by distribution and supply companies, such as Powergen, and by supply companies such as Centrica. Local distribution and supply functions are licensed separately and, in some cases, are separated. The only taboo is on having the high voltage transmission network co-owned with any other activity.

Integration of generation and supply raise no competition issues if neither market exhibits market power, and may offer a means for firms to hedge price variability. In the United Kingdom, the transmission company, the National Grid Company, assists the co-ordination of transmission and generating investment by publishing an annual Seven Year Statement (SYS), which summarises existing demand forecasts made by customers, lists existing and projected generation capacity, and describes the current status of the transmission network and firm future developments[37]. It also indicates 'generation opportunities, defined as areas where new generation can be connected without the need for major transmission reinforcement. By providing such information, the transmission body can further the co-ordination of investment decisions.

These arrangements have made the United Kingdom, with the Netherlands, the poster child of European energy regulation. In the energy review recently completed by the European Commission, other member states have fared less well and many competitive defects have been identified, including extensive vertical foreclosure by integrated companies.

Thus in relation to electricity, the Commission finds (EC (2006) p 169) that:

- vertical integration of generation and retail reduced incentives to trade on the wholesale market, leading to a lack of liquidity in those markets – the resulting volatility of which obliges new entrants to enter in both markets – generation and supply;
- vertical integration of transmission/distribution and supply creates incentives to restrict third party access and to expand the network.

In gas[38]

- access to storage is foreclosed by long-term reservation and investment in new storage capacity may be hampered by the interests of vertically integrated incumbents
- legal and organisational unbundling as foreseen by the current gas directive is not fully implemented. Vertical integration of networks and supply interests leads to conflict of interest resulting in distorted investment incentives.

In the light of this finding the Commission concludes as follows:

> "The Sector Inquiry confirms the finding that it I essential re resolve the systemic conflict of interest inherent in the vertical integration of supply and network activities**,** which has resulted in a lack of investment in infrastructure and in discrimination. It is crucial to ensure that network owners and/or operators do not have incentives that are distorted by supply interests of affiliates. This is particularly important at a time when Europe needs very large investments to ensure security of supply and to create integrated and competitive markets.

[37] See www.nationalgrid.com/uk/Electricity/SYS/
[38] EC (2006) p66.

> Economic evidence shows that full ownership unbundling is the most effective means to ensure choice for energy users and encourage investment. This is because separate network companies are not influenced by overlapping supply/generation interests as regards investment decisions. ...
>
> Furthermore, the public consultation has not revealed any significant synergy effects linked to vertical integration. Indeed, where ownership unbundling has been implemented, experience shows that both the network business and the (production and) supply business continue to thrive after separation."

Energy is an area where a vertical structure causes clear detriments to consumers, often delivered by co-ordinated under-investment. In a system like the United Kingdom's where competitive activities, such as generation and supply, are structurally separated from transmission, indicative forecasting information published by the transmission operator promotes the co-ordination of investment.
Such mechanisms may work equally in other contexts.

4.5. Separation in regulated sectors: an overview

Jose Gomez-Ibanez (2003, pp325-6) has developed a framework, which enables him to give an 'impressionistic assessment' of the costs and benefits of structural separation in a number of sectors. [Recall that the starting point is the imposition of mandatory or regulator-inspired separation.]

The benefits framework starts with the scope for competition, proxied by the share of competitive activities in total costs. Where it is limited, as in the case of supplying small energy and water consumers, the benefits of preventing foreclosure are low. The judgement is also influenced by the scope for innovation in competitive activities: the greater the scope for innovation, the greater the benefit of the competition that separation will bring. Combining these two factors yields an overall benefit evaluation, which, in the case of telecommunications, is high.

The relevant considerations from the cost side are:

- the share of permanent bottleneck assets in total costs, which might be less than the currently non-competitive component; if this is low, it will be relatively inexpensive to build excess capacity in this activity to accommodate any co-ordination failures resulting from separation:
- the degree of heterogeneity of services; the more differentiated the services, the more costly is co-ordination;
- the degree of network interdependence, a higher value of which makes centralised co-ordination more valuable:
- the level of common functions or assets, which are, by definition, difficult or costly to separate.

Gomez-Ibanez estimates that the costs of unbundling telecommunications forms are low. His overall evaluation is shown in the final row of Table 5. The energy and water cases highlight the problems of making a decision where it might have a different effect on small than on large customers. The conclusion on railways reinforces the agnostic or sceptical view on separation noted above. His observations on telecommunications, where he describes the net benefit of separation as 'high', are taken up in Section 5 below.

Table 5. Comparison of the benefits and costs of unbundling across selected industries

	Electricity		Natural gas		Water		Railways		Telecoms
	Small consumer	Large consumer	Small consumer	Large consumer	Small consumer	Large consumer	Freight only	Passenger or mixed	
Factors that affect benefits									
(1) Competitive activities' share of total costs	20-40%	80-90%	20-40%	60-80%	Variable but low	Variable but high	60-80%	50-60%	50-60%
(2) Opportunities for innovation in the competitive activities	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	High
Overall benefit	*Low*	*High*	*Low*	*High*	*Low*	*High*	*High*	*Moderate*	*High*
Factors that affect costs									
(1) Bottleneck infrastructure share of total costs	5-10%	10-20%	20-30%	20-30%	Variable	Variable	20-40%	40-60%	40-50%
(2) Product heterogeneity	Low	Low	Low	Moderate	Low	Moderate	Moderate	High	Moderate
(3) Network interdependence	High	High	Low	Low	Moderate	Moderate	Moderate	Moderate	Moderate
(4) Common functions or assets	Low	Low	Low	Low	Low	Low	Low	Moderate	Moderate
Overall cost	*Low*	*Low*	*Low*	*Low*	*Moderate*	*Moderate*	*Moderate*	*High*	*Low*
Overall advantage	*Low*	*High*	*Low*	*High*	*Low*	*High*	*Low/mod*	*Low*	*High*

Source: Gomez-Ibanez (2003) p328.

B. Unregulated cases

4.6. Vertical integration in the US automobile industry.

The most widely discussed illustration of the harmful effects of separation in an unregulated industry relates to an episode in the US automobile industry in the 1920s. The two companies in question were General Motors and the car body manufacturer Fisher Body.

The case came to prominence because it was employed in an influential article by Klein Crawford and Alchian (2003) -cited in Section 3 above- as an example of the phenomenon of opportunistic behaviour by one firm (Fisher) in a vertical relationship, made possible by the poor specification of the contract. It is alleged that General Motors was forced to resolve the problem by acquiring Fisher Body in 1926.

In particular, it was claimed by the authors that the long-term contract for the supply of bodies entered into gave Fisher the opportunity to recover excessive costs, and that an independent Fisher resisted an efficient location of a new production facility to place it in better position to practice post-contractual opportunism.

However, a string of more recent articles, summarised in Coase (2006), shows that the facts are not as described: the contract did not allow opportunistic behaviour on the scale suggested, and the reason for the vertical integration was quite different than stated above.

This more recent research effectively debunks a mistaken but widely cited argument against the effectiveness of contracts and in favour of integration.

4.7. The personal computer and gaming industry

Personal computers are much like automobiles, they are assembled by companies such as Dell, Gateway, Sony and Toshiba who decide what to make and what to buy. Many components are made by specialised companies, and some are only manufactured by a handful of firms. The vertical chain in personal computing comprises upstream hardware (e.g. central processing units, motherboards, graphics cards, monitors, etc.) and software components (importantly the operating system and BIOS – Basic Input-Output System) and downstream assembly and retailing. A number of well known companies operate in both the upstream and downstream segments, such as Microsoft, Sony and Toshiba.

The main hardware elements produced upstream are the motherboard, which is the circuit board on which other internal elements sit, such as RAM and the CPU. Motherboards are produced by companies such as ASUS, Abit, nVidia, DFI, etc. Motherboards determine what central processor unit (CPU) can be installed into a personal computer or games console. CPUs are manufactured by Intel and AMD for personal computers and by IBM for the main games consoles. Most personal computer motherboards are made to be compatible with either an Intel CPU or an AMD CPU. In January 2007 AMD announced it was developing an open standard specification for a new generation of PC motherboards being more thermally efficient in collaboration with ASUS.[39] This will enable lower noise and more energy efficient PCs. Intel similarly works closely with motherboard manufacturers in producing innovations, and recently Foxconn announced it was going to offer new motherboards integrating graphics cards based on Intel chipsets and nVidia graphic processing units.

[39] http://www.amd.com/us-en/Corporate/VirtualPressRoom/0,,51_104_543~115265,00.html

The personal computing industry is complex and involves close ties between key manufacturers across the transactional boundaries. These collaborative ties are necessary foster and co-ordinate innovation. As personal computers and games consoles comprise many components, these contractual relationships appear a superior means for the OEMs to allow the benefit of competition to flow through innovation.

For example, the Microsoft Xbox 360, a computer console designed for gaming, was developed in co-operation with IBM (who supply the CPU), ATI (who supply the GPU (graphics processor unit)), Samsung, and SiS. Sony produces the main rival to the Xbox 360, the Playstation, and collaborates with IBM and nVidia, the latter being the main rival to ATI in the graphics card market, among others.

No company has chosen to produce all the necessary components in a modern high specification personal computer or games console. However, Sony is more vertically integrated than Microsoft in the gaming console market, producing more of the components and also actively involved in downstream retailing.

Intel and AMD, the main suppliers of central processing units for PCs, have extensive collaborative relationships with manufacturers producing other key components for motherboards. In 2006 AMD announced it was acquiring one of its main collaborators ATI, so that it could develop integrated processing and graphics processors.[40] Interestingly AMD, ATI and its rival nVidia are all under investigation by the Department of Justice in the United States in relation to the pricing of graphics cards.[41]

Intel enjoys a strong or even a dominant position in the manufacture of CPUs for personal computers, and works closely with Microsoft in enabling the development of higher speed processors (for example the Core 2 family of processors were designed to work in compliance with the Windows Vista operating system produced by Microsoft. It also collaborates with other firms in developing and manufacturing complementary products, where it may lack in house expertise or recognises that other firms have superior products. However, given the strength of Intel's position in the processor and related markets, 'smaller' collaborators run the risk that Intel may choose to make rather than buy. Such a perception could damage the long-term interests of companies like Intel and AMD by stifling innovation incentives among smaller collaborators.

To resolve the investment incentives, Intel uses the rhetorical device of distinguishing two Jobs which the company performs: Job 1 - the task of expanding demand for the microprocessor, and Job 2 - the task of growing profitable business in complementary markets. Gawer and Henderson (2007) examine how Intel sought to reconcile conflicts between the two Jobs over the period 1990-2004. It did so essentially by signalling that it would not subsidise entry into complementary markets by driving prices down – by creating separate divisions for Job 2, with profit and loss responsibility. At the same time it committed to not making too much money in these markets by actively giving away intellectual property and subsidising competitive entry. These commitments left it still able to intervene where it had a large comparative advantage.

The key learning here is that firms can manage vertical issues in a highly dynamic complex industry in a sophisticated way, which neutralises the fear of opportunistic behaviour across a transaction boundary.

---

[40] http://www.amd.com/us-en/Corporate/VirtualPressRoom/0,,51_104_543~110899,00.html

[41] http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9005596&intsrc=article_more_bot

4.8. Contracting in the energy sector: power station and take-or-pay gas contracts

Energy assets are typically durable and may be highly specific. Contracts need to be drafted to take these features into account. Two examples are given below which show how this works in practice.

The first example is of a common phenomenon – the simultaneous development of a mine and a minemouth power station. Each side will clearly be concerned about the consequence of post-investment opportunism, and the value of either asset will be lower outside the relationship – though the degree to which this is true will vary from case to case.

In a series of papers, Joskow tests the relations between the characteristics of the parties and the nature of the contracts they draw up.[42] The results show that the more specific the assets, the longer the contract duration. In other words, as repeated *ex post* bargaining becomes more costly and risky, the parties find it worthwhile to spend more on designing and implementing long-term contracts.

The second example relates to the form of contract between natural gas producers and pipelines to which they sell their gas. The producers face the problem that once they have drilled the well, the pipeline companies may put pressure on prices by refusing to take deliveries of the gas. As standard methods of contract enforcement are shown to fail for legal reasons, producers fall back upon the use of take or pay contracts. In one sense, these are inefficient, as they may force pipelines to take gas where they had better alternatives, but this is counterbalanced by their ability to limit opportunism.

An empirical study of contracts (Masten and Crocker 1985) duly shows that the proportion of take-or-pay in the contract was higher the fewer alternative uses there were for the gas producer. In other words, as with power generation contracts, the parties fine tune the provisions to avoid uncertainty.

These are two examples of an extensive contract literature, which cover many sectors[43]. The general result is that contracts can be devised which do take the strain, and that investment is co-ordinated subject to them across contractual boundaries.

5. Implications for telecommunications

5.1. Some general conclusions

It is now time to consider what implications the above analysis has for the structure of an incumbent fixed telecommunications operator in a small country. But it is helpful to start at a broader level of generality.

The theoretical approaches set out above analyse to what degree different activities in a vertical structure are organised hierarchically within a firm and to what extent by market methods involving transactions across boundaries defined by separate ownership. Many of the answers are sought using the concept of transactions costs. Under separation, there are

[42] See Joskow (1987).
[43] See Joskow (2006) and S. Masten (1996).

the so-called 'ink costs' of preparing contracts and also monitoring and enforcement costs. On a wider reading, transactions costs also include the consequences of opportunistic behaviour, which are most acute when assets are specific in terms of their location and use, as well as being durable. In these conditions, one party may be able to exploit the other, and the expectation of this deters investment. Since most telecommunications network assets are highly specific locationally and in respect of their use, this factor is relevant to the present discussion.

However, the case studies above also show how contracting can be developed to deal with these problems, by such means as long term or risk-sharing contracts. We are thus not entitled to say that integration is preferable. It has its costs too, associated with limitations of the span of control and from the fact that different core competencies are needed for different activities. In a situation like this, how should the outcome be determined? Clearly in a competitive market, firms will be forced to seek the structure, which minimises costs. Rational decision making and trial and error on their part is a much better approach than allowing regulators to determine structures.

Now introduce market power in one or more stage of production. If it is just one stage, then, absent regulation, the least-cost organisational form will be chosen, with the bottleneck-controlling firm taking the monopoly profit at that stage, and all other operations being carried out efficiently on either a make or buy basis. If there are two successive stages monopolised by two separate forms, the first will take a monopoly profit or margin and pass it on in higher implicit prices to the second. The result of this double margin will be inefficiently high prices, which can be reduced to the monopoly level by imposing vertical integration.

In practice, this pro-integration result is of almost no relevance because prices set by monopoly or dominant firms' prices will be controlled by regulation. In this context, a contrasting and clear pro-separation result emerges. A vertically integrated firm has the means and motive to practise non-price discrimination in the activity where it is dominant to weaken or eliminate competitors in potentially competitive activities. Recognition of the prevalence of this phenomenon has triggered the current wave of interest in separation.[44]

This provides the basis for a rebuttable prescription in favour of network separation, or unbundling, unless the latter can also be shown to lead to a socially inefficient but privately profitable outcome. This is considered below, but first we recall Gomez-Ibanez's analysis summarised in Table 5 above, where he finds the overall benefit of separation in telecommunications to be high, and the overall costs low. The estimation of benefit is based upon a high proportion of competitive costs (an argument which has to be qualified in the case of small countries such as Ireland), and a lot of scope for innovation in the competitive activities. On the cost side, the relative cheapness of excess capacity in bottleneck elements in telecommunications (compared, for example, with water supply) can be used to mitigate co-ordination problems in the context of network separation: wholesale services are relatively homogeneous, and becoming more so as the industry switches to IP; common assets across either a LoopCo or (*a fortiori*) a NetCo divide are few, and operational network interdependencies can be handled they arise.

On this basis, the author places telecommunications in the category most suitable for unbundling, just as, for broadly opposite reasons, he places railways in the least suitable category. Nonetheless, his evaluation does not seek to focus – as the present one does – on investment issues, and these deserve a fuller discussion.

[44] See section 1.2 above.

5.2. Investment co-ordination in the context of a structurally separated NetCo

In his review of investment incentives under regulation, Guthrie (2006) notes the complexity of the interactions, and the way in which relatively small changes in regulatory rules can affect investment incentives – in most cases, shifting it from one sub-optimal outcome (e.g., under-investment) to another (over-investment). Section 4.1 above provides an illustration of how the type of price cap regulation practised in the UK produces incentives for firms to under-fulfil exaggerated investment plans.

This bias is independent of structure, and it is therefore difficult, as noted above, to distinguish dysfunctional investment outcomes, which result from behavioural regulation from those which result from different organisational structures.

A second key contextual element follows from the earlier general discussion and the case studies in unregulated environments. Those show (and further illustrations can be given) that ingenious contracting forms and sophisticated signalling and commitment behaviour by firms (see the Intel case study in Section 4.7) can act as a substitute for integration in a wide range of cases.

Thirdly, the scale of the problem of investment co-ordination depends crucially on the forms of separation adopted (see Section 2.1). If it is of the LoopCo kind, two problems arise: (i) investment in LoopCo and other (core and backhaul) networks assets have to be co-ordinated across an (ii) ownership boundary which is shifting over time as network architecture changes with the introduction of next generation networks.

Finally, it was noted in Section 2.4 that the co-ordination difficulties of a properly functioning operational separation are virtually identical with those of structural separation, in the sense that allowing investment decision under the former mode to be made at the group rather than the divisional level opens the door to lack of equivalence in the treatment of the affiliated and non-affiliated upstream or downstream organisations. In other words, co-ordination within the hierarchy is inevitably purchased at cost of breaching equivalence.

How then would investment co-ordination problems be resolved in a structurally separated NetCo model, taking as an example the introduction of a step change such as a next generation access (NGA) network. The scale of the problem is in any case somewhat mitigated by the fact that a VDSL model would typically be rolled out progressively, allowing early results to determine the pace of the effort.

There would clearly be an information transmission issue – to ensure that data or expected demand would be received by NetCo. In the first place, NetCo could do the task itself, by consumer research of the type undertaken by most manufacturers of consumer goods, for example. Second, it could have direct (and confidential) communications with retailers, possibly under the supervision of the regulator, in the manner of the constructive engagement between airport operator and airline pioneered by the UK airport regulator (see Section 4.3). Thirdly, NetCo could collate and publish others' forecasts, as is done in the UK electricity supply industry.

There is secondly the risk that NetCo would invest in network assets, while retailers would fail to make the corresponding marketing investments, or would 'hold-up' NetCo by demanding better terms to utilise its new product. Here too experience in other sectors described above comes into play. NetCo could seek to enter into 'take or pay' contracts with retailers, to alter their incentives. In any case, the regulatory system should itself provide and appropriate cost of capital, reflecting both systematic and specific risk (the latter

probably abating over time) as he investment goes forward. The regulator might also be open to representation as Ofcom in the UK is,[45] to incorporate in the cost of capital a real options component to 'buy out' NetCo's tendency otherwise to delay.

By definition a separated structure imposes heavier demands on contracting. But the evidence of studies cited in this paper is that contracting can, in most cases, take the strain. Given that structural separation has one clear advantage in a regulated context -- its ability to drive out anti-competitive conduct, there is no justification for prohibiting it on the basis of theoretical and unsupported conjectures about contracting failures.

[45] Ofcom (2005b).

Bibliography

Averch, H. and Johnson, L. L (1962) 'Behaviour of the firm under regulatory constraint', *American Economic Review*, 52, pp1053-69.

Barzel, Y. (1982) *Measurement costs and the organisation of markets', *Journal of Law and Economics,*25, pp27-48.

CAA (2006)*,Price control review – initial proposals for Heathrow, Gatwick and Stansted Airports*. Civil Aviation Authority.

CAA (2007), *Price control review-CAA recommendations to the Competition Commission for Heathrow, Gatwick and Stansted.* Cave, M.. (2006) 'Six degrees of separation' *Communications and Strategies* No 64, pp89-104*.*

Cave M. et al (2006) Regulating for non-price discrimination-the case of UK fixed telecommunications', *Competition and Regulation in Network Industrie*s, 1, pp391-415.

Competition Commission (2007), *Inquiry into airport changes at Heathrow and Gatwick Airports – issues statement.*

Coase, R. H. (1937) 'The nature of the firm' *Economica*,4, pp386-405.

Coase, R. H. (2006) 'The conduct of economics, the example of Fisher Body and General Motors', *Journal of Economics and Management Strategy*, 15, pp255-278.

Cremer, J., Cremer, K. and De Donder, P. (2006)*, Legal vs. ownership unbundling in network industries*, CEPR Discussion Paper No. 5767.

Crocker, K.J. and Masten, S.E. (1996) 'Regulation and Administered Contracts Revisited: Lessons from Transaction-Cost Economics for Public Utility Regulation', *Journal of Regulatory Economics*, 9, pp5-39.

Crouch, M. (2006) 'Investment under RPI-X: practical experience within incentive-compatible approach in the GB electricity distribution sector' *Utilities Policy*, 14, pp240-4.

Cyrenne, P. (1994) 'Vertical integration vs. vertical separation: an equilibrium model' *Review of Industrial Organisation*, 9, pp11-22.

Department of Transport (2006) *Progress Report.*

Department of Transport (2007*) Delivering a Sustainable Railway*, White Paper Cm7176.

EC (2007*) European Commission Energy Sector Inquiry Report*.

Friebel, G., Ivaldi, M. and Vibes, C. (2004) *Railway (de)regulation: a European efficiency comparison.* CEPR discussion paper 4319.

Gawer A. and Henderson R., (2007) 'Platform owner entry and innovation in complementary markets: evidence from Intel', *Journal of Economics and Management Strategy*, 16, pp1-24.

Gomez-Ibanez, J.(2003) *Regulating infrastructure :monopoly, contracts and discretion,* Harvard University Press.

Guthrie G. (2006) 'Regulating infrastructure: the impact on risk and investment', *Journal of Economic Literature*, 44, pp925-972.

Hart, O. (1995) *Firms, contracts and financial structure*, Oxford University Press.

Joskow, P. (1987) 'Contract duration and relation-specific investments: empirical evidence from coal markets', *American Economic Review*, pp168-185.

Joskow, P. (2006) *Vertical integration*, mimeo.

Klein, B., Crawford, RA and Alchian, A.A.(1978) 'Vertical integration, appropriable rents and the competitive contracting process'*, Journal of Law and Economics*, 21, pp297-326.

Laffont, J-J and Tirole, J. (1993) *A theory of incentives in procurement and regulation*, MIT Press.

Masten, S. (1996) *Case Studies in Contracting and Organisation*, Oxford University Press.

Masten S. and Crocker, K.(1985), 'Efficient adaptation in long-term contracts, take-or-pay provisions for natural gas, *American Economic Review*, pp1083-1093.

MED (2007) , *Development of Requirements for the Operational Separation of Telecom: Consultation Document*, Ministry of Economic Development, New Zealand.

Milgrom, P. and Roberts, J., (1992) *Economics organisation and management*, Prentice Hall.

OECD (2001) *Recommendation of the Council concerning structural separation in regulated industries.*

OECD (2003*) The benefits and costs of structural separation in the local loop,* DSTI/ICCP/TISP(2002)13/FINAL.

OECD (2005) *Structural reform in the railway industry* DAF/COMP (2005) 46, OECD, 2005.

OECD (2006*) Report of experiences in the implementation of the Recommendation concerning structural separation in regulated industries.*

Ofcom (2005a) *A Notice under Section 155(1)of the Enterprise Act 2002.*

Ofcom (2005b*) The cost of capital: statement.*

Ofcom (2006*) Survey of BT's wholesale customers.*

Office of Fair Trading (2006) *UK Airports.*

Pittman, R. (2004) 'Structural Separation and access pricing in the railways section: sauce for the goose only?' pp. 35-47 of *Structural Separation and Access Pricing: a New Synthesis?*

TCNZ (2007a) *Submission of Telecom New Zealand limited in response to MED consultation document*

TCNZ (2007b) *Development of requirements for the operational separation of Telecom,* TCNZ.

Temin, P and Galambos, L. (1987) *The Fall of the Bell System*, Cambridge University Press.

Williamson O. (1985) *The Economic Institutions of Capitalism*, Free Press.