University of Leicester

DEPARTMENT OF ECONOMICS

TOC 'N' ROLL: BARGAINING, SERVICE QUALITY AND SPECIFICITY IN THE UK RAILWAY NETWORK

Gianni De Fraja, University of Leicester, Università di Roma "Tor Vergata" and C.E.P.R.

Emanuela Michetti, University of Siena, Italy

Piercarlo Zanchettin, University of Leicester, UK

Working Paper No. 09/07
February 2009

Toc 'n' Roll: Bargaining, Service Quality and Specificity in the UK Railway Network*

Gianni De Fraja†
University of Leicester, Università di Roma "Tor Vergata" and C.E.P.R.

Emanuela Michetti‡ — University of Siena

Piercarlo Zanchettin§ — University of Leicester

February 27, 2009

Abstract

The paper studies the regulatory design in an industry where the regulated downstream provider of services to final consumers purchases the necessary inputs from an upstream supplier. The model is closely inspired by the UK regulatory mechanism for the railway network. Its philosophy is one of vertical separation between ownership and operation of the rolling stock: the Train Operating Company (TOC) leases from a ROlling Stock COmpany (ROSCO) the trains it uses in its franchise. This, we show, increases the flexibility and competitiveness of the network. On the other hand, it also reduces the specificity of the rolling stock, thus increasing the cost of running the service, and the TOC's incentive to exert quality enhancing effort, thus reducing the utility of the final users. Our simple model shows that the UK regime of separation may in fact be preferable from a welfare viewpoint.

JEL Numbers: D2; L1; L5; L92

Keywords: Network regulation, Railways, Incomplete contracts, Relation specific investment.

---

*We wish to thank audiences at the Italian Society in Law and Economics and in Sheffield and Newcastle.

†University of Leicester, Department of Economics Leicester, LE1 7RH, UK, Università di Roma "Tor Vergata", Dip. SEFEMEQ, Via Columbia 2, I-00133 Rome, Italy, and C.E.P.R., 90-98 Goswell Street, London EC1V 7DB, UK; email: `defraja@le.ac.uk`.

‡Department of Economics, University of Siena, Piazza San Francesco 7, I-53100, Siena, Italy; email: `emanuela.michetti@unisi.it`.

§Department of Economics, University of Leicester, University Road, LE1 7RH, Leicester, UK; email: `pz11@le.ac.uk`.

1 Introduction

We study the interaction between two regulated firms: a downstream provider of services to final consumers and the upstream supplier of the inputs necessary for the provision of the services. While the analytical set-up is closely modelled on the scheme chosen in the UK for the regulation of the railway network and therefore it can yield immediate practical interpretations for that case, it also has broader lessons for the regulatory design of the contractual relationship between upstream and downstream firms in regulated industries.

The defining regulatory feature of the UK regulatory system[1] in place for the railway network is the separation between the Train Operating Company (TOC) and the ROlling Stock COmpany (ROSCO). In a nutshell, the national network is divided into several geographical franchises,[2] essentially separate from an operational viewpoint. In each of these areas, one TOC provides the rail services to passengers, and leases its rolling stock from one of the three ROSCOs which are allowed to operate. Each franchise is assigned to a TOC by the Department for Transport, following a competitive process.[3] The aspect of the process relevant to our paper is that, to be allowed to bid, a TOC must enter a broad agreement for the supply of the necessary rolling stock with one of the ROSCOs. Once a TOC is awarded the franchise for one area, it is essentially constrained to lease the rolling stock it needs from the specific ROSCO it had chosen before bidding. Following the award of the franchise, the original broad agreement between them can be "finalised", or, in the economic jargon, re-negotiated. We pick up the thread from here, in a model which studies the re-negotiation between the two parties and the effects of the agreement they reach. In line with the UK regulatory mechanism, we rule out the possibility for the TOC to switch to a different ROSCO.

The UK regulatory mechanism is intended to enhance competition. This can happen through three channels. First, the well understood competitive effect of bidding for the award of the franchise is strengthened, in the UK

[1] An exhaustive and up-to-date description of the regulatory and institutional set-up of the UK railway industry is in Office of Rail Regulation (2007); see Newbery (1997) for an overview of the utilities' reform in the UK.

[2] Currently, there are 19 franchised operators, the Department for Transport is the awarding body for all of them except three. Additionally, there are 6 non-franchised operators. A full list of the train operators in the UK can be found at: www.atoc-comms.org/franchised-passenger-services.php.

[3] See ORR (2007) for details on the franchise process. For a general overview on ORR's role after the implementation of the Railways Act 2005, see Department for Transport (2007).

mechanism, by the fact that participation in the bidding does not require a TOC to incur the hefty sunk entry cost constituted by the acquisition of suitable rolling stock: this increases the number of potential bidders. Secondly, precisely because the incumbent does not own the rolling stock, it is feasible to award the franchise for a period shorter than the rolling stock's economic life,[4] ensuring that the threat to withdraw the franchise is credible and keeps the incumbent on its toes. Finally, preventing the TOCs from owning rolling stock makes it easier to transfer rolling material from one area to another, which may be required following changes in the geographical distribution of demand.

The ease with which rolling stock can be switched to a different franchise is a key determinant of the efficacy of the latter two channels in enhancing competition, and therefore we build a model that centres around it. Many technical factors influence how easily rolling stock can be switched. The type of the train's power supply, its maximum speed, the position its doors, its size, clearance and configuration (a train designed with a flat rural region in mind, with few large stations, may be totally unsuited to a mountainous area with many smaller stations).[5] If some of these technical factors, such as the gauge, are effectively fixed, most can instead be typically varied by whoever designs the train. In other words, the *specificity* of the rolling stock, which affects how easy it is to deploy a train to an area different from the one it was designed for, is chosen to balance technical considerations with economic and strategic reasons. Our model studies the trade-off between increasing specificity, that is designing the rolling stock in a way suited to the geographical area to which it is destined, and decreasing specificity, that is increasing flexibility by opting for a design which makes it easier to operate the rolling stock in a different area. We argue here that the regulatory system itself affects the choice of the level of specificity of the rolling stock, and, through it, the quality and cost of the train services.

A suitable conceptual framework to study the role played by the regulatory regime on the interaction between TOC and ROSCO is provided by the theoretical literature on incomplete contracts.[6] This is based on the idea that

[4] "Engines and carriages have a working life far longer than the length of a passenger franchise contract, and are therefore not owned by the companies themselves but by private sector leasing companies.", UK Department for Transport (2004).

[5] See SRA (2004) for further details about the degree of standardisation in the rail network in the UK.

[6] Hart and Holmstrom (1987) provided an early review, and Tirole (1999) an evaluation.

in many long term relationships, a party who can make a relation specific investment which reduces the costs and/or increases the benefits of the other party, may refrain from doing so if it is unable to reap a share of the benefit of its investment. This happens because contracts are incomplete, in the sense that it is impossible (or prohibitively costly) to specify the obligation of each party in every conceivable eventuality in sufficient detail to allow a third party, called to enforce the contract in the event of a dispute, to determine whether a breach has occurred or not.[7] We cast the choice of the train specificity s as a relation specific investment by the ROSCO. The investment in our paper has however a conceptually different nature from the investment in the incomplete contracts literature. The cost of specificity is not given by the production process, as there is no reason to suppose that building a "flexible" train is in principle more or less costly than building a highly specific train, but instead by the lower net revenues which can be obtained using the rolling stock in a different region. This is the nature of specificity: more specificity helps produce a high quality service on the "right line", but it decreases the quality – and hence the market value – of a train's services on the "wrong line".[8] This has the subtle implication that the *cost* of the "investment in specificity" depends on the regulatory mechanism. The barriers to writing a complete contract are determined by the UK regulatory regime, which imposes the separation between TOC and ROSCO. In this sense, the degree of separation between the TOC and the ROSCO must itself be seen a policy instrument, and the paper provides a conceptual framework to analyse its role and its effects. Our set-up, therefore, differs from the standard incomplete contract literature: the feasible contracts are not exogenously given by technological and informational constraints, and so it makes sense to compare the "separation" regime chosen for the UK rail network, akin to contract incompleteness between the TOC and the ROSCO, with the "integration" regime of complete contracts, typical of most other EU countries.[9] We compare two cases. In the first, complete

[7] The typical example is a clause specifying that quality must be "good" or "adequate": in the event of a dispute, even though both parties may come to the same (private) judgement as to whether quality is "adequate", an enforcer, such as a court or an arbitrator, cannot.

[8] See Ménard and Yvrande-Billon (2005) for a similar point of view: "The non-redeployability is critical here. Discrepancies between contract duration and the physical lifetime of equipment exist in many leasing industries (e.g. car and truck rental), but it is no problem as long as equipment has alternative users".

[9] The regulatory regime is however clearly under strain, to the point that the Office of Rail Regulation has recently referred to the competition authority about the prevention of

contracts can be written and the TOC and the ROSCO agree *both* on transfer prices *and* on the design of the train. In the second, the parties are not permitted to write complete long term contracts which specify in sufficient detail the characteristics of the rolling material to be supplied by the ROSCO to the TOC, and they can only agree on transfer prices, while the degree of specificity of the rolling stock is chosen by the ROSCO on its own.

We find that both the degree of specificity and the investment in quality increase with integration, in line with most of the literature.[10] The fact that specificity and quality increase with integration does not however necessarily imply that socially a fully integrated structure should be preferable, as suggested, among others, by Ménard and Yvrande-Billon (2005) and Preston (2002). Indeed, our model shows that there can be *over*-investment in specific assets and *excessive* quality of service. Too much specificity may mean too little competition as the TOC and the ROSCO become too closely locked together and sheltered from competition for the franchise: the technological benefit of specificity is traded off the lack of flexibility and the anticompetitive effect of highly specialised rolling stock. This trade-off implies that a case by case analysis is in principle necessary to evaluate the best regulatory design.

The paper is organised as follows: Section 2 introduces the model: demand, technology and the possible vertical structures in 2.1, and the bargaining mechanism in 2.2. The temporal sequence of events and decisions is summarised in Section 2.3. The policy analysis begins in Section 3 with the determination of the first best choice of specificity and effort; these are compared in Section 5 with the equilibrium values in the two regulatory regimes derived in detail in Section 4. The proofs of the more algebraic results are in the Appendix, which is preceded by a brief conclusion.

2 The Model

2.1 Demand, technology, and regulatory regimes

We model the interaction between three agents: a regulator, the firm franchised to supply rail transport services in a given region, train operating company, or

competition in the leasing market of rolling stock for franchised passenger services. In August 2008 the Competition Commission published its provisional findings and confirmed that some of the features of the rolling stock leasing market do indeed raise competition issues. See ORR (2007) and Competition Commission (2008).

[10] See Kain (1998), Preston (2002), Crompton and Jupe (2003), Ménard and Yvrande-Billon (2005).

TOC in what follows, and the firm who has the expertise to design and supply rolling stock, trains, and locomotives, ROSCO hereafter.

Consumers care about the quality of the service they receive.[11] This is denoted by q, and depends on the investment of the TOC, $e \in [0,1]$ (a normalisation),[12] and on the realisation of a random variable, θ, with cumulative distribution function $\Phi(\theta)$, density $\phi(\theta) = \Phi'(\theta)$, and support also normalised to $[0,1]$. For concreteness, in the case of the railways, we can think of quality q as given by the frequency of delayed trains, of the TOC's investment, e, as its provision and arrangements with regard to stand-by personnel and equipment, and of the random element, θ, as the external exogenous factors that affect the provision of the train services. A high investment in stand-by provision by the TOC reduces the negative impact, in terms of inconvenience and delays, of a negative quality shock, such as a derailment.[13] For simplicity we take an additive specification:

$$q = \theta + e, \tag{1}$$

and assume $\Phi(\theta)$ to be uniform:

$$\Phi(\theta) = \theta. \tag{2}$$

Total revenues for rail service in the region are exogenously given by $r > 0$. Although quality can vary, revenues are independent of quality. This may happen, for example, because the marginal consumer does not value quality, even though the inframarginal consumers do, or because the price is regulated and the regulator does not use mechanisms (such as the one studied by De Fraja and Iozzi 2007) linking the allowed prices to the realised quality. This simplification helps us concentrate on the relationship between quality and specificity, leaving aside the interaction between prices and quality.

Once the market structure is established, the regulator's tool-kit is the imposition of sanctions in the event of deterioration of the quality of the service. We capture these sanctions with the simplifying assumption that the regulator

[11] Consumer surplus and total surplus are formalized in Section 3.

[12] Constraining e not to exceed 1 should be seen as capturing the idea that the cost of effort increases very rapidly when e approaches the technological maximum.

[13] θ are adverse weather shocks and freak accidents (such as the disaster caused by a SUV becoming stuck on the railway and derailing the Newcastle-London high speed train near Selby on 28/2/2001). e are the measures taken by companies to minimise the probability of accidents and reducing the disruption caused by the weather (rather than attributing delays and cancellations to leaves on the line or to the wrong "type of snow", as famously commented by a British Rail executive on 11/2/1991).

chooses a minimum quality requirement $q_m \in [0, 2]$ and the TOC's franchise is renewed if and only if its service quality is at least the minimum requirement, $q \geqslant q_m$. While most sanctions take the form of a fine, the draconian punishment of withdrawing the licence was imposed on Connex South Eastern, a train operator serving the South-East of England, in the second half of 2003 (National Audit Office, 2005).[14] Given (1) and (2), the franchise is renewed with probability

$$z = 1 - q_m + e. \tag{3}$$

By investing in quality enhancing activities e, the TOC can increase the quality of the service and hence reduce the probability that the franchise is not renewed by the regulator.

The ROSCO's investment is the degree of train specificity, denoted by s. Without loss of generality, we also normalise it to lie in $[0, 1]$.

Specificity has benefits and costs. On the one hand, it reduces the cost of providing quality, see (5) below. On the other hand, it makes it more costly to transfer rolling material to a different area: the more technically and operationally suitable a train is to network A, the less suitable it is to network B, and so we posit that the net revenue that can be obtained from using it on network B is a decreasing function of the degree of specificity to the original TOC. Formally, we denote by $P(s)$, with $P'(s) < 0$, the unit net revenues that can be obtained using a train of specificity $s \in [0, 1]$ destined to a network different from the one it was designed for. We also assume that $P(s)$ is *concave* in s: this would follow, for example, from the natural assumption of convex adjustment costs. We specify $P(s)$ as:

$$P(s) = P_0 - \frac{1}{2}ks^2, \tag{4}$$

where $k > 0$, and $P_0 \geqslant \frac{1}{2}k$, so that the revenue from an alternative network are non-negative for every value of s. Note that there is in general no reason to presume that specificity has also a technological cost: a train with, say fixed height entry steps (specific to the design of the stations where the TOC operates) need not be more expensive to design and build than a train with variable height entry steps, which can be transferred to rail networks with a different station design.

[14] An alternative, analytically identical, assumption is to posit some uncertainty on the TOC's part with regard to the regulator's preference: the regulator is satisfied and renews the franchise with a probability which is increasing in the quality level offered by the TOC.

The TOC and the ROSCO are private profit maximising companies. As explained above, the TOC's revenues are exogenously given by $r > 0$. The TOC incurs two types of costs: the payments to the ROSCO for the use of its rolling material, endogenously determined and discussed in details in the next section, and the operating costs. The latter depend positively on the TOC's effort for quality, since quality is costly, and negatively on the specificity of the rolling material supplied by the ROSCO, since specifically designed trains have lower running costs. We assume a linear specification:

$$C(e, s) = c_0 + c_e e - c_s s, \tag{5}$$

where $c_0, c_e,\ c_s > 0$, and $c_0 > c_s$: the last ensures that operating costs are positive for every possible combination of e and s. We deliberately set the cross derivative, $\frac{\partial^2 C(\cdot)}{\partial e \partial s}$ at 0, to isolate the interaction between e and s which is caused by the institutional set-up from any technological complementarity.

While exogenously given and independent of quality, demand is not fixed. It may be adversely affected by an idiosyncratic shock, such as the closure of a local employer or the opening of a new motorway, which affects demand in one area, but not in the others. Without shock, total revenues are $r > 0$. The negative shock reduces demand by a proportion $u \in (0, 1)$, and happens with probability $(1 - x)$. This makes $\alpha = u\,(1 - x)$ the relevant measure of the TOC's *expected* loss in the demand for the final service.[15] We eliminate the uninteresting possibility that the train service is shut down by assuming that the line is profitable even if the demand shock occurs: $ur - c_0 > c_e$.

We assume that the areas of the network are symmetric, which imposes the analytically convenient restriction that the unit profit obtained from employing a generic train (i.e., a train with $s = 0$) in an alternative region is the same as the unit profit from using a generic train in the area considered:

$$P_0 = r - c_0 - c_e e_A, \tag{6}$$

where e_A is the investment in effort chosen by the passengers service operator in the alternative area. Clearly, $ur - c_0 > c_e$ implies $r - c_0 - c_e > 0$, so that

[15] It is convenient to rule out aggregate uncertainty: the number of areas affected by the negative demand shock is known in advance, as is which franchisees will be able to lease the additional trains from the areas where the adverse demand shock has occurred: the only uncertainty is which areas will receive the adverse demand shock. The exact formalisation of this simplification would require distinguishing the analysis of TOCs who will be able to lease some of their trains from the adversely affected areas, from those who will not: this would add heavy and unrewarding notation.

the industry profit from running generic trains is always positive across the network.

The salient feature of the regulatory regime of the UK rail industry is the separation between the ROSCO, the firm who owns the trains, and the TOC, the firm who runs them, enshrined in the ban of long term agreements for the supply of the rolling stocks. We translate this legal requirement into the assumption that s is chosen separately by the ROSCO, which does so with the aim to maximise its own profit. If instead the TOC and the ROSCO were allowed to negotiate s, then they would choose it to maximise their joint profit, and so their choice of s would be identical to that made by continental Europe style integrated company.[16] We refer to the UK regulatory regime as the "vertical separation" regime, and to the joint negotiation of s as the "vertical integration" regime.

2.2 Bargaining

The ROSCO can be involved in three bargaining situations. We model them all as generalised Nash bargaining, with exogenously given bargaining power coefficient β for the ROSCO. In this model, the values agreed upon in the bargaining process are those which maximise the weighted sum of the log of the parties' surplus, with weights β for the ROSCO and $(1 - \beta)$ for its counterpart (the well known details are spelled out in Appendix 1).

The first bargaining situation for the ROSCO is its negotiation with the TOC over the lease contract for the provision of the rolling stock and, in the vertical integration regime, also over the specificity of the rolling stock. We assume that the lease price negotiated by the TOC and the ROSCO is a two-part tariff (p, F), where F is a fixed fee, and p is the unit price of the train services actually leased. p is only paid if the TOC's franchise is renewed, and varies with demand. In our set-up, choosing p and F is equivalent to choosing e, the quality-enhancing effort by the TOC, and therefore our assumption of bargaining over a two part tariff corresponds to bargaining over e. As a consequence, for a given degree of train specificity, the effort for service quality chosen by the TOC always maximizes the joint profit generated by the TOC and ROSCO relationship irrespective of whether they are vertically integrated

[16] Whether the two companies are integrated or legally separated entities is not relevant to the choice of the train design: if the companies are legally separated they will also negotiate a side payment, which depends on the relative bargaining power of the two parties, and affects the distribution, but not the size, of the total profit.

or separated.[17] To determine the generalised bargaining solution it is necessary to know the "disagreement payoff"of the two parties, which in the model is their outside option. The TOC's is 0: in the event of disagreement between the TOC and the ROSCO, the TOC loses the franchise, which is reassigned to an alternative TOC. We label the latter "ATOC", an acronym reminiscent of the actual Association of TOCs. On the other hand, the ROSCO's outside option is in general strictly positive because it can lease the rolling stock to the ATOC following a broken down negotiation; it is also endogenously determined.

The second bargaining situation in which the ROSCO can be involved is the negotiation with the ATOC over the rolling stock to be used in the franchise under consideration. This can happen for two reasons: either because the TOC and the ROSCO do not reach an agreement in their negotiation or because the TOC loses its franchise due to its service quality falling short of the minimum standard set by the regulator. In the first case, the ROSCO and the ATOC bargain on a lease contract for the provision of the rolling stock and, in the vertical integration regime, on the specificity the rolling stock. In the second case, the ROSCO and the ATOC bargain only on the lease contract even in the vertical integration regime, since the train specificity has been irreversibly set in the previous agreement between the ROSCO and the TOC. In both situations, the bargaining between the ROSCO and the ATOC takes place after the ATOC has already set its investment in effort for quality and after the quality uncertainty has been resolved, but before the resolution of the uncertainty in demand. In other words, in both situations the regulator reassigns the franchise to an "average" ATOC which has already passed the quality control in another area of the network. Bargaining on a linear price is not distortionary, since the effort for quality is given, and both parties have zero outside options, as a disagreement would lead to the cancellation of the service.[18]

The ROSCO's third possible bargaining situation is triggered by the ad-

[17] This modelling strategy is justified since information asymmetries between the TOC and the ROSCO regarding e are probably not fundamental (one presumes that this was the underlying assumption on which the separation imposed by the UK legislator is based), and the aim of the paper is to concentrate on the comparison between the regulatory regimes, not on the role of information asymmetries.

[18] That is, there can be at most one "negotiation breakdown". The analysis would be unaffected if we assumed instead that there is a (finite) sequence of potential ATOC's with which the ROSCO could negotiate. We would simply need to work backwards from the last bargaining process in the sequence, adding complication but no insight.

verse demand shock occurring in the area served by the TOC. The TOC needs fewer trains, and the ROSCO offers the surplus rolling stock to the operator of train service in a different area. In this case, the ROSCO and this "external" operator bargain when all the relevant variables (the train specificity, the "external" operator's investment in effort, and the random shocks) are fixed and known to the two parties. As before, bargaining over a unit price p is not distortionary, and both parties have a zero outside option in this case.

The bargaining power coefficient satisfies:

$$\frac{\beta}{(1-\beta)^2} > \frac{1}{2k}\frac{c_s^2}{c_e}\frac{1-\alpha}{\alpha}. \qquad (7)$$

This guarantees that relevant second order conditions are satisfied. To ensure symmetry across the network, all exogenously fixed parameters are the same in all areas of the network.

2.3 Timing

The timing of choices in multi-stage games affects the outcome of the interaction among players. The timing is determined by constraints imposed by the regulatory regime and by technology, and, in our set-up, the investment required to design and build rolling stock has clearly a longer time span than the investment in quality enhancing effort, implying that the model must be such that s is chosen before e. Demand and quality shocks have a relatively short term nature and therefore can be posited to occur after s and e have been set. The parties operate in a fixed regulatory regime, that is, they know that the regulator's rules and guidelines regarding the link between the minimum quality standard and the likelihood of sanctions being imposed will not be changed as a consequence of the parties' actions.[19]

These considerations lead to the following formal description of the timing of the game.

1. **Regulatory set-up.** The regulator chooses whether the TOC and the ROSCO will negotiate over the triple (s, p, F) – the vertical *integration regime* – or over (p, F) only – the *vertical separation* regime.

[19]It is of course possible that the regulatory standard is unexpectedly tightened after the parties actions, for example as a consequence of a media campaign: conceptually, this would correspond to a negative quality shock. We restrict the range of combinations of causal effects with the assumption that the regulatory standard is not affected by the choice of the parties.

2. **Minimum quality.** The regulator selects q_m, the minimum quality necessary for the franchise to be continued.

3. **Train specificity.** The train specificity, s, is decided by the ROSCO in the vertical separation regime, and negotiated by the TOC and the ROSCO in the vertical integration regime.

4. **Lease contract.** The TOC and the ROSCO bargain over (p, F) , the terms under which the ROSCO's rolling material can be used by the TOC.

5. **Effort.** The TOC decides its effort for quality, e.

6. **Quality uncertainty resolution.** The quality shock is realised.
   (a) If the realised service quality offered by the TOC is at least q_m, the TOC retains the franchise and operates the service in the area.
   (b) If the realised service quality offered by the TOC is below q_m, then the TOC loses the franchise, which is reassigned to the ATOC. The ROSCO and the ATOC bargain on the leasing price, p (because the train specificity and ATOC's effort for quality have already been determined in previous stages). If they reach an agreement, the ATOC operates the service in the area, otherwise the service is cancelled.
   (c) Similarly, if the TOC and the ROSCO have failed to reach an agreement in their Stage 4 negotiation, the franchise is cancelled and reassigned to the ATOC. In this case, the ROSCO and the ATOC negotiate over the leasing price, p, and, in the vertical integration regime, also over the rolling stock specificity, s. Like in stage 6(b), if they reach an agreement, then the ATOC operates the service in the area, otherwise the service is cancelled.

7. **Demand uncertainty resolution and payoffs.** The demand shock is realised.
   (a) If the service is operated in the area under consideration, the realized demand is served by the TOC (or by the ATOC), which leases the rolling stock it needs from the ROSCO, pays the ROSCO according to contract and collects passengers revenues.

(b) If, due to adverse demand conditions, there are surplus trains in the area, the ROSCO negotiates with the franchisee of a different region. If they agree, the ROSCO collects a share β of the unit profit $P(s)$ times the quantity of train services transferred to the "external" franchisee. In the case of disagreement, the unused rolling stock remains idle, and the ROSCO makes no profit from it.

3 Social welfare and first best

The aim of the paper is the comparison of the performance of different regimes against the yardstick of industry social welfare, measured by the sum of the expected consumers' and producers' surplus. In this section, we characterise the benchmark given by the first best social optimum. We consider a representative area of the network, and invoke symmetry to extend our findings to the rest of the industry.

The producers' surplus is the profit which can be obtained from a non-specific train, *viz* a train with $s = 0$, $r - c_0 + c_e e$, augmented by the expected cost saving due to the train specificity to the area under consideration, $(1 - \alpha) c_s s$, and reduced by the expected extra cost due to the train specificity when the train is used in a different area, $\alpha \frac{1}{2} k s^2$. We assume that the consumers' surplus depends only on the final service quality, q, not on the train specificity s. We choose again a convenient functional form and let consumers' surplus be given by:

$$\sigma q \left(1 - \frac{b}{4} q\right).$$

$b \in (0, 1]$ and $\sigma > 0$ measure the concavity of the consumers' welfare function, and the intensity of consumers' preference for quality. σ could also be interpreted as the importance of consumers' surplus relative to profit. The restriction $b \in (0, 1]$ ensures that consumers' surplus increases with q in its range, the interval $[0, 2]$. Recalling that $q = e + \theta$ and θ is uniformly distributed on $[0, 1]$, we can write the consumers' surplus as a function of e only, say $S(e)$:

$$S(e) = \sigma \frac{(6 - b) + 3(4 - b)e - 3be^2}{12}.$$

The expected social welfare, $W(e, s)$, is given by:

$$W(e, s) = (r - c_0 - c_e e) + (1 - \alpha) c_s s - \frac{\alpha k}{2} s^2 + \sigma \frac{(6 - b) + 3(4 - b)e - 3be^2}{12}. \tag{8}$$

The first best benchmark is the choice of investment in quality, e, and of train specificity, s, which maximizes (8). At an interior solution these are:

$$s^* = \frac{(1-\alpha)\, c_s}{\alpha k}, \tag{9}$$

$$e^* = \frac{4\sigma - 4c_e - \sigma b}{2\sigma b}. \tag{10}$$

From (9) we can see that the first best degree of train specificity, s^*, increases with its effectiveness in reducing the operating costs on the "right line", c_s, and the expected quantity of train services to be employed on that line, $(1-\alpha)$, and decreases with the importance of the extra costs for using the rolling stock on "wrong lines", k, and the expected quantity of train services to be moved to those lines, α. On the other hand, (10) shows that the optimal level of the investment in quality, e^*, increases with the intensity of consumers' preference for quality (or, in the alternative interpretation, with the importance of consumers' surplus relative to profit), σ, while it decreases with the sensitivity of the operating costs to the effort for quality, c_e, and the concavity of the consumers' welfare function, b.

Graphically, the social welfare function (8) generates in the (s, e)-plane a map of elliptic iso-welfare curves centred at the first-best point (s^*, e^*), as illustrated below in Figure 1.

4 Industry equilibrium

Keeping (9) and (10) as benchmarks, we can now characterise the subgame perfect equilibrium of the game constructed in Section 2 for a representative franchise, working backward from the last decision stage. By symmetry, since all franchises are alike, it constitutes the industry equilibrium.

4.1 Expected profits (stage 6)

Under both vertical regimes, if the game reaches stage 6(b) or stage 6(c), if, that is, the TOC's franchise is not renewed because of an adverse quality shock or because negotiations fail, then the ROSCO negotiate with ATOC a lease price p for the use of the rolling stock available. Since both parties have zero outside option, and since the price does not affect any subsequent decision, the bargained price will distribute the joint profit, $(r - c_0 - c_e e_A + c_s s)$, which is fixed, according to the bargaining power coefficients, β and $(1-\beta)$. Before the resolution of demand uncertainty, the expected quantity of train services to be employed on the line equals $(1-\alpha)$. Therefore, the ROSCO's expected profit from leasing the rolling stock to the ATOC is:

$$(1-\alpha)\,\beta\,(r-c_0-c_e e_A+c_s s)\,.$$

If the TOC loses the franchise because its service quality falls short the minimum standard, stage 6(b), then the train specificity is fixed in both vertical regimes, and the revenues generated by the use of the rolling stock outside the area are unrelated to the outcome of the negotiation between the ROSCO and the ATOC. In the vertical separation regime, the same is true when the game reaches stage 6(c), that is when the TOC loses its franchise because of a lack of agreement with the ROSCO at stage 4. In all these cases, the ROSCO's expected profit from reaching an agreement with the ATOC is:

$$\pi_R^{O_S}=(1-\alpha)\,\beta\,(r-c_0-c_e e_A+c_s s)+\alpha\beta P(s), \tag{11}$$

where s is the rolling stock specificity set at stage 3, and $\alpha\beta P(s)$ is the ROSCO's expected profit from leasing the expected quantity α of unused rolling stock to an external franchisee. $\pi_R^{O_S}$ in (11) is therefore the ROSCO's outside option in the negotiation with TOC under vertical separation.

On the other hand, if the TOC and the ROSCO do not reach an agreement in the vertical integration regime, in stage 6(c) the ROSCO and the ATOC bargain over both prices and the train specificity. The bargained value of s in this case maximises the expected joint profit of the ROSCO and the ATOC, $(1-\alpha)\,(r-c_0-c_e e_A+c_s s)+\beta\alpha P(s)$, and the ROSCO would keep a share β of this maximized profit. Its payoff would therefore be:

$$\pi_R^{O_I}=(1-\alpha)\,\beta\,(r-c_0-c_e e_A+c_s s)+\beta^2\alpha P(s). \tag{12}$$

Analogously to (11), (12) is the ROSCO's outside option in the initial negotiation with the TOC under the vertical integration regime.[20]

We can now write the expected profits of the TOC, Π_T, and the ROSCO, Π_R, following an agreement in their negotiations (stages 3 and 4) and after

[20] As shown below, in the vertical integration equilibrium, the TOC and the ROSCO agree on the rolling stock specificity that maximizes their expected joint profit (given by equation (15) below). Therefore, the outside option $\pi_R^{O_I}$ does not affect the equilibrium values of the rolling stock specificity and the service quality: it will affect only the division of the expected joint profit between the two parties.

the TOC's choice of the effort for quality (stage 5), but before the resolution of quality uncertainty (stage 6) and of demand uncertainty (stage 7):

$$\Pi_T = (1-\alpha)(1-q_m+e)(r-c_0-c_e e+c_s s-p)-F, \tag{13}$$

$$\begin{aligned}\Pi_R = &(1-\alpha)\left[(1-q_m+e)p+\right.\\ &\left.+\,(q_m-e)\beta(r-c_0-c_e e_A+c_s s)\right]+\alpha\beta P(s)+F,\end{aligned} \tag{14}$$

where $P(s)$ is obtained from (4) and (6) as $(r-c_0-c_e e_A)-\frac{1}{2}ks^2$.

In (13), the expected quantity of train services employed on the line, $(1-\alpha)$, and the probability that the TOC's retains the franchise, $(1-q_m+e)$, multiply the difference between total revenue, r, and total costs. The latter is the sum of production costs, $c_0+c_e e-c_s s$, and the unit price paid for the lease of the rolling stock, p. In addition, the TOC pays the fixed fee F to the ROSCO. This appears with a positive sign in the ROSCO's profit, (14). In addition to it, the ROSCO's payoff is the weighted average, with weights α and $1-\alpha$, of (i) the ROSCO's expected revenues per unit of train service operated on the line: with probability $(1-q_m+e)$, the TOC operates the service and the ROSCO collects revenue p; with probability (q_m-e), on the contrary, the ATOC operates the service and the ROSCO collects revenue $\beta(r-c_0-c_e e_A+c_s s)$; and (ii) the ROSCO's expected profit from leasing its unused trains to an "external" franchisee, as discussed in Section 2.2.

(13) and (14) give the expected joint profit generated by the TOC and ROSCO relationship, $\Pi_J \equiv \Pi_T+\Pi_R$:

$$\begin{aligned}\Pi_J = &(1-\alpha)\left[(1-q_m+e)(r-c_0-c_e e+c_s s)+\right.\\ &\left.(q_m-e)\beta(r-c_0-c_e e_A+c_s s)\right]+\alpha\beta P(s).\end{aligned} \tag{15}$$

4.2 Effort for service quality (stage 5)

At stage 5, the TOC takes the train specificity and the terms of the leasing contract determined in (13) as fixed, and chooses e to maximise Π_T. The first order condition $\frac{\partial \Pi_T}{\partial e}=0$ is:

$$(r-c_0-c_e e+c_s s-p)-(1-q_m+e)c_e=0.$$

Hence the TOC's profit maximising effort choice is:[21]

$$e_5 = \frac{r-c_0+c_s s-p-c_e(1-q_m)}{2c_e}, \tag{16}$$

provided this is in $[0,1]$.

[21] The second order condition for an interior solution requires $-2c_e<0$, which is satisfied.

4.3 Lease contract (stage 4)

(15) and (16) appear to suggest that the unit price p affects the Stage 5 choice of e by the TOC, and therefore, on top of its redistributive effect, it also exerts a distortionary effect. However, as we show in Appendix 1, this distortionary effect disappears when the parties negotiate over the fixed fee F, as they do here. To see this intuitively, let Π_R^O denote the ROSCO's outside option in the case of disagreement.[22] Irrespective of how the rolling stock specificity s is chosen, the two parties set their contractual terms, (p, F), in such a way that the unit-price p induces the TOC to choose the joint profit maximising level of effort e_5:

$$\frac{\partial \Pi_J}{\partial e} \frac{\partial e_5}{\partial p} = 0. \tag{17}$$

The fixed fee F distributes the surplus generated by the relationship, $\Pi_J - \Pi_R^O$, according to the bargaining power coefficients β and $1 - \beta$:

$$\Pi_R = \Pi_R^O + \beta(\Pi_J - \Pi_R^O), \tag{18}$$

$$\Pi_T = (1 - \beta)(\Pi_J - \Pi_R^O). \tag{19}$$

From (16), $\frac{\partial e_5}{\partial p} = -(2c_e)^{-1} < 0$, and, from the first order condition of the TOC's optimisation problem at stage 5, $\frac{\partial \Pi_T}{\partial e} = 0$. Hence, condition (17) reduces to $\frac{\partial \Pi_R}{\partial e} = 0$, where Π_R is given in (14). This leads to:

$$p_4 = \beta \left(r - c_0 - c_e e_A + c_s s \right). \tag{20}$$

From equations (16) and (20) we can now derive the equilibrium level of the effort for quality, as a function of the rolling stock specificity:[23]

$$e_4 = \frac{(1 - \beta) \left(r - c_0 + c_s s \right) + \beta c_e e_A - c_e \left(1 - q_m \right)}{2c_e}. \tag{21}$$

In the symmetric equilibrium of the industry, the rolling stock specificity and the effort for quality are the same in all areas of the network, so that $e_A = e_4$, and (21) gives the following expression for the equilibrium level of the effort for quality, $\widehat{e}$, as a function of the equilibrium level of the rolling stock specificity, $\widehat{s}$ ($\widehat{e}$ and $\widehat{s}$ are common to all franchises):

[22] As shown above, the ROSCO's outside option in the negotiation with the TOC depends on the vertical regime: it is given by equation (11) under vertical separation and by equation (12) under vertical integration. This difference is, however, immaterial for our argument here.

[23] In alternative, solving in e the condition $\frac{\partial \Pi_J}{\partial e} = 0$, and using the solution in system with (16), would lead the same expressions for e_4 and p_4 as in equations (21) and (20), respectively.

$$\widehat{e} = \frac{(1-\beta)(r - c_0 + c_s \widehat{s}) - c_e(1-q_m)}{c_e(2-\beta)}. \tag{22}$$

4.4 Choice of specificity (stage 3)

In the vertical integration regime, the rolling stock specificity is part of the negotiation between the TOC and the ROSCO.[24] In equilibrium, the TOC and the ROSCO agree on the level of specificity which maximises their joint profit Π_J, given in (15) (details are in Appendix 1). The first order condition $\frac{\partial \Pi_J}{\partial s} = 0$ yields:[25]

$$\beta + (1-\beta)(1 - q_m + e) = \beta \frac{\alpha k}{(1-\alpha) c_s} s. \tag{23}$$

In the vertical separation regime, the rolling stock specificity, s, is unilaterally set by the ROSCO, to maximize its expected profit which would follow an agreement with the TOC:

$$\max_s \ \Pi_R = \Pi_R^{O_S} + \beta(\Pi_J - \Pi_R^{O_S}), \tag{24}$$

where Π_J and $\Pi_R^{O_S}$ are given by equations (15) and (11), respectively (see again Appendix 1). Notice that the ROSCO's outside option, $\Pi_R^{O_S}$, itself depends on the rolling stock specificity.[26] Using (15) and (11), the first order condition $\frac{\partial \Pi_R}{\partial s} = 0$ yields:

$$1 + (1-\beta)(1 - q_m + e) = \frac{\alpha k}{(1-\alpha) c_s} s. \tag{25}$$

We have assumed that costs are linear and parameters are constrained to lie within an interval. This is an approximation of the observation that, in

[24] Note that, while s and (p, F) are chosen in separate stages, under the vertical regime, the "decision maker" is the same in the two stages, and sequential or simultaneous solution yield the same outcome. Formally, let $M(s, p, F)$ be the joint payoff function of the TOC and the ROSCO. In simultaneous decision making, the optimal (s, p, F) satisfy the first order conditions: $\frac{\partial M}{\partial s} = \frac{\partial M}{\partial p} = \frac{\partial M}{\partial F} = 0$. In sequential decision making the parties take s as given when choosing (p, F), and so: they set $\frac{\partial M}{\partial p} = \frac{\partial M}{\partial F} = 0$. This gives p and F as functions of s, say $\tilde{p}(s)$ and $\tilde{F}(s)$. In the previous stage, they set s, taking into account the effect of their choice on their own future choices: the first order condition is: $\frac{dM}{ds} \equiv \frac{\partial M}{\partial s} + \frac{\partial M}{\partial p}\tilde{p}'(s) + \frac{\partial M}{\partial F}\tilde{F}'(s) = 0$. But of course $\frac{\partial M}{\partial p} = \frac{\partial M}{\partial F} = 0$, and $\frac{dM}{ds} = 0$ reduces to $\frac{\partial M}{\partial s} = 0$.

[25] The second order condition for an interior solution is satisfied by (7).

[26] Since $\Pi_R^{O_S}$ does not depend on e and $\frac{\partial \Pi_J}{\partial e} = 0$, any indirect effect of s on Π_R via e vanishes in the first order condition $\frac{\partial \Pi_R}{\partial s} = 0$. The second order condition for an internal equilibrium requires: $\frac{\alpha k}{(1-\alpha)c_s} - \frac{1}{2}\frac{c_s}{c_e}(1-\beta)^2 > 0$, which is implied by (7)

practice, costs raise progressively more rapidly as one variable moves towards more extreme values, and simplifies the analysis conveniently. The price paid for this simplification is the need to consider many corner cases at stage 3, since, for some values of the parameters, the equilibrium values of some variables are at an end point of the interval.

Proceeding with the analysis, consider first the case $s \in (0,1)$. To compact notation, in what follows we denote:

$$K \equiv \frac{\alpha k}{(1-\alpha)\, c_s}, \qquad R \equiv \frac{r - c_0}{c_e} > 1, \qquad C \equiv \frac{c_s}{c_e}, \tag{26}$$

and:

$$\delta = \begin{cases} \beta & \text{for the integration regime} \\ 1 & \text{for the separation regime} \end{cases}.$$

δ is a parameter characterising the vertical regime. The following values of the variables characterise a "fully interior equilibrium" – an equilibrium where s, e and z (the probability that TOC's franchise is continued, given in (3)) all lie in the interior of their respective range – and are obtained by solving the first order conditions (23) and (25) to determine the equilibrium values of s, and then substituting in (22) and (3) to obtain e and z:

$$\hat{s} = \frac{(1-\beta)^2\,[R + (1-q_m)] + \delta\,(2-\beta)}{\delta\,(2-\beta)\,K - C\,(1-\beta)^2}, \tag{27}$$

$$\hat{e} = \frac{\delta\,(1-\beta)\,(RK + C) - (1-q_m)[\delta K - C\,(1-\beta)^2]}{\delta\,(2-\beta)\,K - C\,(1-\beta)^2}, \tag{28}$$

$$\hat{z} = \frac{\delta\,(1-\beta)\,[K(1-q_m) + RK + C]}{\delta\,(2-\beta)\,K - C\,(1-\beta)^2}. \tag{29}$$

If the equilibrium is not characterised by the first order conditions, it is given by values of the choice variables at the boundary of the choice set. The equilibrium set is fully described in Lemma 1. Let:

$$\Gamma = \delta\,(1-\beta)\,[(R-1)K + C] - [\delta K - C\,(1-\beta)^2]. \tag{30}$$

Lemma 1 *There exist threshold values, $q_1^z, q_1^e, q_0^e \in (0,2)$, such that:*

i) If $\Gamma \geqslant 0$:

$q_m \in [0,1]$	*implies*	$e = q_m$	$s = \frac{1+\delta-\beta}{\delta K}$	*and* $z = 1$,
$q_m \in [1,2]$	*implies*	$e = 1$	$s = \frac{\delta+(1-\beta)(2-q_m)}{\delta K}$	*and* $z = 2 - q_m$.

ii) If $0 \geqslant \Gamma \geqslant -(1-\beta)\,\delta K$*:*

$$\begin{array}{lllll} q_m \in [0, q_1^z] & \textit{implies} & e = q_m & s = \frac{1+\delta-\beta}{\delta K} & \textit{and } z = 1, \\ q_m \in [q_1^z, q_1^e] & \textit{implies} & e = \hat{e} & s = \hat{s} & \textit{and } z = \hat{z}, \\ q_m \in [q_1^e, 2] & \textit{implies} & e = 1 & s = \frac{\delta+(1-\beta)(2-q_m)}{\delta K} & \textit{and } z = 2 - q_m. \end{array}$$

iii) If $-(1-\beta)\,\delta K \geqslant \Gamma$*:*

$$\begin{array}{lllll} q_m \in [0, q_0^e] & \textit{implies} & e = 0 & s = \frac{\delta+(1-\beta)(1-q_m)}{\delta K} & \textit{and } z = 1 - q_m, \\ q_m \in [q_0^e, q_1^e] & \textit{implies} & e = \hat{e} & s = \hat{s} & \textit{and } z = \hat{z}, \\ q_m \in [q_1^e, 2] & \textit{implies} & e = 1 & s = \frac{\delta+(1-\beta)(2-q_m)}{\delta K} & \textit{and } z = 2 - q_m. \end{array}$$

Moreover, the threshold values are given by:

$$q_1^z = 1 + \frac{\delta\,(1-\beta)\,[(R-1)K + C] - [\delta K - C\,(1-\beta)^2]}{\delta\,(1-\beta)\,K}, \tag{31}$$

$$q_1^e = 2 - \frac{\delta\,(1-\beta)\,[(R-1)K + C]}{\delta K - C\,(1-\beta)^2}, \tag{32}$$

$$q_0^e = 1 - \frac{\delta\,(1-\beta)\,(RK + C)}{\delta K - C\,(1-\beta)^2}. \tag{33}$$

The gloriously tedious proof is in Appendix 2. Lemma 1 describes how the equilibrium changes according to the value of the quality standard q_m. In particular, for low standard, that is if $q_m \leq q_1^z$, the TOC's franchise is renewed with probability 1, to the point that, for $q_m \leq q_0^e$ the equilibrium level of the effort for quality becomes 0. Conversely, this effort is at its maximum value, $e = 1$ for q_m high enough, that is when $q_m \geq q_1^e$.

5 Optimal regulation

When the regulator cannot directly choose the first best values of the rolling stock specificity and the effort for service quality, she needs to influence them indirectly by setting the minimum quality standard, q_m. The regulator takes as given the vertical regime – separation or integration – and chooses the quality standard that maximises the social welfare function (8) under the constraint that the values of the rolling stock specificity and the effort for quality be the equilibrium values for that regime.

We begin the analysis of the regulator's policy by showing that, in both vertical regimes, she faces a trade-off between specificity and effort for quality.

Definition 1 *The* regulation possibility locus *is the locus of points in the (s, e)-plane representing the combinations of effort for quality and rolling stock specificity achievable through the regulation of the minimum quality standard.*

While the regulator's choice of q_m can affect the interaction between TOC and ROSCO, and their choice of e and s, only combinations (s, e) on the regulation possibility locus can be induced by the choice of q_m. These are described in next result.

Proposition 1 *In both vertical regimes, the regulation possibility locus is non-increasing, and it is strictly decreasing when the equilibrium values of s, e and z all lie in the interior of their respective ranges.*

Proof. For all cases where the equilibrium is not "fully interior", Lemma 1 shows that, as q_m increases, either the effort for quality increases while the rolling stock specificity remains constant, or the rolling stock specificity decreases while the effort for quality remains constant. Consider now fully interior equilibria. Solving (27) for q_m and substituting in (28), yields:

$$\widehat{e} = R + \frac{\delta}{(1-\beta)^2} - \frac{1}{(1-\beta)^2}\left[\delta K - C(1-\beta)^2\right]\widehat{s}, \tag{34}$$

that is, the relation between $\widehat{e}$ and $\widehat{s}$ is linear. Next, according to Lemma 1, a fully interior equilibrium requires $\Gamma < 0$, which clearly implies $\delta K - C(1-\beta)^2 > 0$. Therefore, (34) is a downward sloping line in the (s, e)-plane. Finally, it is clear from (27) and (28) that an increase in q_m corresponds to a movement along this line where $\widehat{s}$ falls and $\widehat{e}$ rises. ■

Figure 1 illustrates two possible shapes of the regulation possibility loci in the separation regime (the solid line) and in the integration regime (the dotted line). The LHS diagram illustrates the second part of Lemma 1. For low values of q_m, TOC's franchise is renewed with probability 1, and the loci are vertical: as q_m increases, e increases but s remains constant. For intermediate values of q_m, the equilibrium is fully interior, and the loci take the shape of a downward sloping line. For high values of q_m, the effort for quality is at its maximum possible level and the rolling stock specificity and the probability that TOC renews its franchise both decrease with q_m: the loci become horizontal on the $e = 1$ line. The picture illustrating the first part of Lemma 1 is different from the diagram on the LHS only in that the downward sloping portion of the curve is vertical. The RHS diagram depicts the third part of Lemma 1. When q_m is low, the effort for quality takes its minimum possible value, $e = 0$, with z and

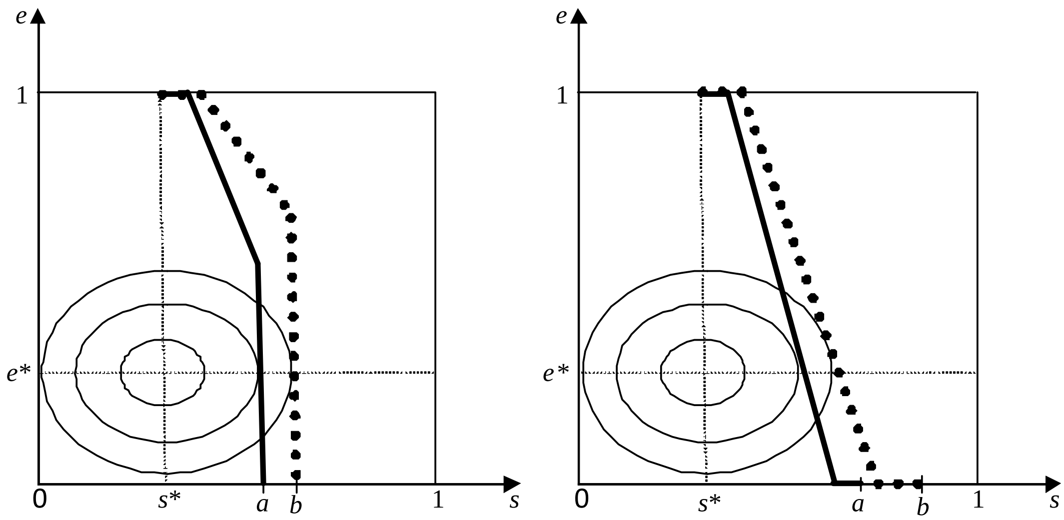

Figure 1: The iso-welfare loci and the regulation possibility loci.

s still in the interior of their respective ranges, and the loci exhibit a horizontal portion on the e-axis. For intermediate values of q_m, the equilibrium is "fully interior" and the two loci are downward sloping. For high values of q_m, the effort for quality is fixed at its maximum possible level: this is the horizontal portion of the two loci at $e = 1$.

The intuition behind Proposition 1 is the following. As the regulator raises the quality standard, the probability that TOC's franchise is renewed decreases, even though the TOC partly offset the higher quality threshold with its effort for quality. This reduces the private expected return of specificity in both regimes. This is because whoever chooses the level of s (the ROSCO on its own or the TOC and the ROSCO in concert) has a lower marginal return from specificity if the ATOC rather than the TOC operates the service. Under vertical integration, the reason is that the cost saving due to more specificity is fully internalized in the TOC and ROSCO joint profit when the TOC operates the service, while a share $(1 - \beta)$ is captured by the ATOC if the TOC's franchise is not renewed. Under vertical separation, the ROSCO has a positive outside option in the negotiation with the TOC which increases with the rolling stock specificity. Therefore, an increase in specificity raises the ROSCO's share of the joint profit realised with the TOC.

Our next result shows that, irrespective of the vertical regime, the industry equilibrium always exhibits over-investment in rolling stock specificity unless the regulator sets q_m at its maximum value, $q_m = 2$, in which case that TOC's franchise is renewed with probability 0.

Proposition 2 *Let* $0 < \frac{c_s(1-\alpha)}{k\alpha} < 1$ *and* $\beta < 1$*. Then, in both vertical regimes,* $q_m < 2$ *implies* $\widehat{s} > s^*$*.*

Proof. Using the notation introduced in (26), we first rewrite the first order conditions (23) and (25) as:

$$\delta + (1-\beta)\widehat{z} = \delta \frac{\alpha k}{(1-\alpha)\, c_s} \widehat{s},$$

where $\widehat{z}$ is the equilibrium probability that TOC's franchise is continued. Since $\frac{\alpha k}{(1-\alpha)c_s} = \frac{1}{s^*}$ (see equation (9)), we have:

$$\widehat{s} = s^* \left[1 + \frac{(1-\beta)}{\delta} \widehat{z} \right],$$

which ensures that, in both regimes, there is over-investment in specificity unless $\beta = 1$ or $\widehat{z} = 0$. Finally, by Lemma 1, $\widehat{z} = 0$ if and only if $q_m = 2$. ■

The first condition in the statement ensures that the first-best socially optimal level of rolling stock specificity is interior, $0 < s^* < 1$, and the second that the ROSCO does not have full bargaining power.

Proposition 2 is illustrated in Figure 1. All possible cases formalised in Lemma 1 generate second-best loci with the qualitative characteristic depicted in the figure: the dotted line is (weakly) to the right of the solid line, and both are to the right of the first best (s^*, e^*): irrespective of the vertical regime, to achieve any value of $e < 1$, the regulator must accept over-investment in specificity. Only by setting $q_m = 2$ the regulator can eliminate the over-investment in s, and implement the first-best socially optimal level of specificity, with the resulting value of e at its maximum possible level.

In a word, Proposition 2 says that there is over-investment in specificity. To understand why this happens, recall that specificity offers the social benefit of lowering the cost of operating the rolling stock on the network it was designed for, at the social cost of additional operating costs if the train is moved to a different network. The private choice of s would replicate the first-best socially optimal choice only when the social cost and the social benefit of specificity are internalised in the profit function of the private decision maker exactly in the same proportion. This is clearly the case when the ROSCO has full bargaining power, in which case the social cost and benefit of specificity are fully internalised in the ROSCO's profit in both vertical regimes. If instead the ROSCO does not have full bargaining power, if, that is, $\beta < 1$, then only a share β of the social cost of specificity is internalised in the profit function of the decision maker for the choice of s, the remaining $1 - \beta$ share is borne by an ATOC operating the service in a different network. On the benefit side, if the TOC's franchise is not renewed ($z = 0$), then the TOC and the ROSCO (in the integration regime) or the ROSCO alone (in the separation regime) collect the fraction β of the profit generated by the ATOC on the "right" network, and thus the share β of the social benefit of specificity: the social cost and benefit of specificity are internalised in the same proportion by the private decision maker, and the private choice of s is socially optimal. If, on the contrary, $z > 0$, then, irrespective of the vertical regime, the private decision maker of s internalises (in expectation) a share of the social benefit of specificity larger than β:[27] the private decision maker of s internalises a higher proportion of the social benefit than of the social cost of specificity, which leads to over-investment in specificity in both regimes.

We are now in the position to analyze the regulator's optimal choice of the vertical regime.

Proposition 3 *If the first-best social optimum is interior in both s and e, the vertical separation regime is strictly socially preferable to the vertical integration regime.*

The "interior" condition implies that the centre point of the ellipses drawn in the picture is not on the sides of the $[0, 1] \times [0, 1]$ square in the (s, e)-plane.

The regulator selects the minimum quality standard to obtain the combination of e and s given by the point of tangency of the separation locus and the highest possible iso-welfare curve. As the Figure illustrates, the reason why the vertical separation regime (solid line) is preferable is that it limits the over-investment in specificity. While the intuition is easily illustrated, the formal proof is more complex, and can be found in Appendix 2. Notice also that, irrespective of the shape of the regulation possibility loci, the regulator's

[27] As argued in the discussion of Proposition 1, the private marginal return of specificity is always higher when the TOC rather than the ATOC operates the service on the "right" network, and it equals the fraction β of social marginal benefit when the ATOC operates the service.

preferred value of e is 1 only when the first best is also 1. In this case, the optimal regulation policy achieves the first-best levels of both s and e by setting $q_m = 2$ in both regimes.

The intuition for Proposition 3 is straightforward given our discussion of Propositions 1 and 2. The over-investment in specificity is always stronger under vertical integration: if the TOC operates the service on the "right line", the social benefit of specificity is fully appropriated by the ROSCO and the TOC in the vertical integration regime, whilst only a fraction of it is appropriated by the ROSCO in the vertical separation regime. On the other hand, as we have already explained, in both regimes whoever sets s internalises the fraction β of the social benefit of specificity when the ATOC operate the service on the "right line", and the fraction β of the social cost of specificity in all circumstances.

In the second-best solution, there is always over-investment in specificity (unless the regulator can achieve the first-best), while the level of the effort for quality can be greater or equal to the first-best value. This is shown in Figure 1. If the tangency point between the separation locus and the lowest possible iso-welfare curve is on a downward sloping (respectively, vertical) portion of the locus – as on the RHS (respectively, LHS) –, then the second best is a fully interior equilibrium and there is (respectively, there is no) over-investment also in effort for quality. Notice also that the regulation of the quality standard, linked to the threat of franchise termination, is an effective way to fine tune the effort for quality, which can take any value in $[0, 1]$, and the regulator adjusts the investment in effort for quality with the aim to reduce the over-investment in specificity.[28]

[28] We have so far assumed that the equilibrium level of specificity is interior: $\widehat{s} \in (0, 1)$. This is not restrictive. We have already shown that $\widehat{s} \geq s^*$, and therefore the second best optimum, $\widehat{s}$, can be 0 only when the first best, s^*, is also 0, that is when specificity has no social value. This is not realistic for the railway industry. $\widehat{s} < 1$ is not restrictive either, provided the first best level, s^*, is also lower than 1. Allowing for $\widehat{s} = 1$ simply means that our second best loci now can reach the line where $s = 1$, and partially coincide with it. However, for all other portions of the loci, our comparative results would still apply. This means that the integration locus will never lie below the separation locus, and the second best solution under integration never dominates the second best solution under separation. More precisely, the two regimes would offer the same social value: (i) for $e^* = 1$ (as shown in our second-best analysis), since the over-investment in specificity can be completely eliminated without any cost in term of over-investment in effort; (ii) for sufficiently low values of e^*, since, for these low values of e, both regimes would offer the same level of specificity, the upper bound $s = 1$. In all other cases (i.e., for intermediate values of e^*) vertical separation dominates vertical

6 Concluding remarks.

This paper examines the effects of imposing separation on the vertically related suppliers of the outputs necessary in a regulated industry. The stylised model is inspired by the structure of the UK railway industry, where TOCs and ROSCOs, the suppliers of train services and of rolling stock are obliged to maintain a substantial degree of separation from each other. This is unlike most of the rest of the world, where instead the suppliers of train services also own the trains used to supply those services. It can shed light on several other regulated industries, where it is technologically feasible to separate the vertical stages of production, in some cases, such as British Gas, imposed as a remedy to non-competitive practices.

Our paper confirms the view that the UK system provides a weaker incentive to specific train design and effort for quality than a more vertically integrated system. However, unlike the existing literature (see footnote 10 above), we show that the benefit of a competitive and flexible structure, both ex-ante and ex-post, may well outweigh the negative effects of lower specificity and the lower quality effort that occur in the separation regime, as also suggested by Affuso and Newbery (2002).

In the competitive environment which characterises the UK railways, specificity is set partly with the aim of reducing the probability that the franchise is lost by the Train Operating Company: a high specificity makes it less costly for the TOC to meet the quality standard set by the regulator. To the extent that the ROSCO can extract some of the surplus from the TOC being abler to comply with the quality standard, the ROSCO can gain by increasing the rolling stock specificity even though it will mean lower revenues in the event that trains need to be switched to a different area. And here lies the difference between the two vertical regimes. In the integration regime, the ROSCO can extract more of the TOC's surplus, because it can do so in the direct negotiation. In the separation regime, it must instead do it indirectly, by adjusting s itself, which affects the ROSCO's outside option in the negotiation with the TOC over the lease contract. The surplus it can extract is smaller, and hence the downside of specificity (the lower revenues in the case specific trains need to be switched to a different line) keeps its level down in the separation regime.

integration for the same reasons illustrated in this Section. Finally, the case where $s^* = 1$ is a limit case where the two regimes are always equivalent. This shows that the regime of integration is weakly dominated by the regime of vertical separation.

In other words, train specificity is used by the TOC (possibly in concert with the ROSCO) to blunt the regulatory threat of the withdrawal of the franchise (by making it cheaper for the TOC to meet the service quality standard). This over-investment in specificity is greater in the integration regime, because the TOC and the ROSCO are better able to agree on this strategic use of s, and it disappears when it is not possible to affect the probability that the TOC's franchise is continued, that is when, although the equilibrium probability that the franchise is renewed is 0, the effort for quality is already at its maximum value. On the other hand, the equilibrium probability that the TOC's franchise is continued can be 1 only if the TOC adjusts the effort for quality one-to-one to any increase in the quality standard. The cost savings from more specificity play a crucial role in incentivising the TOC to follow this strategy. This explain why, in both regimes, the over-investment in specificity is maximum when $z = 1$.

References

Affuso L. and Newbery D. (2002): The Impact of Structural and Contractual Arrangements on a Vertically Separated Railway, *Economic and Social Review*, 33 (1), pp. 83–92.

Competition Commission (2008): *Rolling Stock Leasing Market Investigation – Notice of Possible Remedies Under Rule 11 of the Competition Commission Rules and Procedures*, London: Competition Commission.

Crompton G. and Jupe R. (2003): A Lot of Friction at the Interfaces: The Regulation of Britain's Privatised Railway System, *Financial Accountability and Management*, 19 (4), pp. 397–418.

De Fraja G. and Iozzi A. (2008): The Quest for Quality: A Quality Adjusted Dynamic Regulatory Mechanism, *Journal of Economics and Management Strategy*, forthcoming.

Hart O. and Holmstrom B. (1987): The Theory of Contracts, *Advances in Economic Theory, Fifth World Congress*, edited by Truman Bewley, Cambridge: Cambridge University Press.

Kain P. (1998): The Reform of Rail Transport in Great Britain, *Journal of Transport Economics and Policy*, 32 (2), pp. 247–266.

Ménard C. and Yvrande-Billon A. (2005): Institutional Constraints and Organisational Changes: the Case of British Rail Reform, *Journal of Economic Behavior and Organisation*, 56 (4), pp. 675–699.

National Audit Office (2005): *The South Eastern Passenger Rail Franchise*, House of Commons, Hc 457.

Newbery D. (1997): Privatisation and Liberalisation of Network Utilities, *European Economic Review*, 41 (3), pp. 357–383.

Office of Rail Regulation (2007): *The Leasing of Rolling Stock for Franchised Passenger Services. ORR's Reasons for Making a Market Investigation Reference to the Competition Commission*; Office of Rail Regulation.

Preston J. (2002): The Transaction Cost Economics of Railways, *Trasporti Europei*, 7 (20-21), pp. 6–15.

Strategic Rail Authority (2004): *Improving Passenger Rail Services through New Trains*, House of Commons, Hc 408.

Tirole J. (1999): Incomplete Contracts: Where Do We Stand?, *Econometrica*, 67 (4), pp. 741– 781.

UK Department for Transport (2007): *Annual Report 2007*, Cm 7095.

UK Department for Transport (2004): *White Paper - The Future of Transport: A Network for 2030*, Cm 6234.

Appendices. Not for publication.

Appendix 1.

In this appendix, we first show that the generalized Nash bargaining between the ROSCO and the TOC over the two part pricing schedule (p, F) leads, as claimed in the text, to the choice of the joint profit maximising level of effort for quality and to the division of the resulting surplus, $\Pi_J - \Pi_R^O$, according to the bargaining power coefficients β and $(1-\beta)$. Subsequently, we show that the rolling stock specificity bargained by the TOC and ROSCO in the vertical integration regime maximizes their joint profit.

In both vertical regime, the TOC and the ROSCO set p and F to solve the maximization problem:

$$\underset{p,F}{Max} \left(\Pi_R - \Pi_R^O\right)^{\beta} \Pi_T^{1-\beta}, \tag{A1.1}$$

where Π_R and Π_T are given by equations (14) and (13), respectively, and Π_R^O is the ROSCO's outside option in the negotiation with TOC. For the moment, we do not need to distinguish between the values of the ROSCO's outside option in the two vertical regimes, but only to note that the ROSCO's outside option is independent of the pricing schedule under negotiation with the TOC. The first order conditions for problem (A1.1) yield:

$$\beta \left(\Pi_R - \Pi_R^O\right)^{\beta-1} \Pi_T^{1-\beta} \frac{\partial \Pi_R}{\partial F} + (1-\beta) \left(\Pi_R - \Pi_R^O\right)^{\beta} \Pi_T^{-\beta} \frac{\partial \Pi_T}{\partial F} = 0$$
$$\beta \left(\Pi_R - \Pi_R^O\right)^{\beta-1} \Pi_T^{1-\beta} \frac{\partial \Pi_R}{\partial p} + (1-\beta) \left(\Pi_R - \Pi_R^O\right)^{\beta} \Pi_T^{-\beta} \frac{\partial \Pi_T}{\partial p} = 0,$$

or

$$\beta \Pi_T \frac{\partial \Pi_R}{\partial F} + (1-\beta) \left(\Pi_R - \Pi_R^O\right) \frac{\partial \Pi_T}{dF} = 0 \tag{A1.2}$$
$$\beta \Pi_T \frac{\partial \Pi_R}{\partial p} + (1-\beta) \left(\Pi_R - \Pi_R^O\right) \frac{\partial \Pi_T}{\partial p} = 0. \tag{A1.3}$$

Now, from equations (14) and (13), we calculate $\frac{d\Pi_R}{dF} = -\frac{d\Pi_T}{dF} = 1$, which means that F has a pure distributive role with no effect on the joint profit. Hence (A1.2) reduces to

$$\beta \Pi_T = (1-\beta) \left(\Pi_R - \Pi_R^O\right), \tag{A1.4}$$

The joint profit is $\Pi_J = \Pi_T + \Pi_R$, and so (A1.4) gives the distribution of the surplus stated in equation (18):

$$\Pi_R = \Pi_R^O + \beta \left(\Pi_J - \Pi_R^O\right),$$
$$\Pi_T = (1-\beta) \left(\Pi_J - \Pi_R^O\right).$$

Next, (A1.4) and (A1.3) imply:

$$\frac{\partial \Pi_R}{\partial p} + \frac{\partial \Pi_T}{\partial p} = 0,$$

that is,

$$\frac{\partial \Pi_J}{\partial p} = 0. \tag{A1.5}$$

Obviously, the pure distributive effect of the unit-price p on the TOC's and the ROSCO's profits cancels out in the joint profit. Therefore, Π_J depends on p only through the effort for quality chosen by TOC at the subsequent stage 5, and (A1.5) is equivalent to:

$$\frac{\partial \Pi_J}{\partial e}\frac{\partial e_5}{\partial p} = 0,$$

which is (17) in the text.

Consider now the vertical integration regime. The ROSCO and TOC bargain over the pricing schedule and the rolling stock specificity simultaneously. Their maximisation problem becomes:

$$\underset{p,F,s}{Max}\left(\Pi_R - \Pi_R^{O_I}\right)^{\beta}\Pi_T^{1-\beta}, \tag{A1.6}$$

where the ROSCO's outside option, $\Pi_R^{O_I}$, is given by (12). Notice that, beside being independent of the pricing schedule, $\Pi_R^{O_I}$ is also independent of the rolling stock specificity under negotiation between the ROSCO and the TOC, since, in the case of disagreement, the rolling stock specificity will remain indeterminate and will enter the successive negotiation between the ROSCO and the ATOC. In addition to (A1.2) and (A1.3), the optimisation problem (A1.6) has the following first order condition in s:

$$\beta\Pi_T\frac{\partial \Pi_R}{\partial s} + (1-\beta)\left(\Pi_R - \Pi_R^{O}\right)\frac{\partial \Pi_T}{\partial s} = 0, \tag{A1.7}$$

which, using (A1.4), becomes:

$$\frac{\partial \Pi_R}{\partial s} + \frac{\partial \Pi_T}{\partial s} = 0.$$

In other words, the TOC and the ROSCO agree on the specificity level that maximizes their joint profit.

Appendix 2.

Proof of Lemma 1. Part *(i)*. We first prove that if $\Gamma \geqslant 0$, (Γ is given in (30)), then there cannot be a fully interior equilibria for any $q_m \in [0,2]$. To this end, reformulate equation (28) as:

$$\widehat{e} = 1 + \frac{\delta(1-\beta)((R-1)K+C) - (\delta K - C(1-\beta)^2) - (1-q_m)(\delta K - C(1-\beta)^2)}{\delta(2-\beta)K - C(1-\beta)^2}. \tag{A2.1}$$

Recall also that:

$$\delta(1-\beta)[(R-1)K+C] > 0 \;\; \text{(since } R>1 \text{ and } C>0\text{)}, \tag{A2.2}$$

$$\beta(2-\beta)K - C(1-\beta)^2 > 0 \;\; \text{(for } \widehat{s} \text{ to be positive and finite)}. \tag{A2.3}$$

If $\Gamma \geqslant 0$, $\widehat{e} < 1$ would require $(\delta K - C(1-\beta)^2) > 0$ and $q_m < 1$, as it is apparent from (A2.1). On the other hand, (31) would imply $q_1^{\tilde{z}} > 1$. Since $\widehat{z}$ is decreasing in q_m (see (29)), $\widehat{z}$ is corner at 1 for $q_m \leq q_1^{\tilde{z}}$. With $q_1^{\tilde{z}} > 1$, $\widehat{z}$ will be corner at 1 for any $q_m \leq 1$. Then, over the $[0, 2]$ range of q_m, the the equilibrium is either corner in e (for $q_m \in (1, 2]$), or corner in z (for $0 \leq q_m \in [0, 1)$), or corner in both e and z (for $q_m = 1$), and we never find a fully interior equilibrium.

To characterise the equilibrium, we first re-write the first order conditions in s of the two regimes (equations (23) and (25)) in the compact form:

$$\delta + (1-\beta)z = \delta K s. \tag{A2.4}$$

For $q_m \in [0, 1]$, $z = 1$, so that $z = 1 - q_m + e$ implies $e = q_m$. The equilibrium level of s must solve condition (A2.4) with $z = 1$, yielding: $s = \frac{1+\delta-\beta}{\delta K}$.

For $q_m \in [1, 2]$, $e = 1$ so that $z = 1 - q_m + e$ implies $z = 2 - q_m$. The equilibrium level of s must solve condition (A2.4) with $z = 2 - q_m$, yielding: $s = \frac{\delta+(1-\beta)(2-q_m)}{\delta K}$.

Parts *(ii)* and *(iii)*. We first prove that, if $\Gamma < 0$, then there exist a sub-interval of $[0, 2]$ such that if q_m is in that sub-interval, then the industry equilibrium is fully interior.

Notice that, $\Gamma < 0$ and (A2.2) imply $\delta K - C(1-\beta)^2 > 0$. Then, from (A2.1), $\widehat{e}$ is increasing in q_m.

Next, $\Gamma < 0$ implies $1 < q_1^e < 2$ (from (32)) and $0 < q_0^e < 1$ (from (33)). This ensures that $\widehat{e}$ takes interior values for values of q_m in a connected sub-interval of $[0, 2]$ containing 1. On the other hand, $\Gamma < 0$ implies $q_1^{\tilde{z}} < 1$ (from (31)). Since $\widehat{z}$ is decreasing in q_m and cannot be zero for $0 < \widehat{e} < 1$ (as explained in footnote 28, p.19), $\widehat{z}$ also takes interior values in a (connected) interval of q_m around 1. There must therefore exist a connected interval of q_m around 1 where both $\widehat{z}$ and $\widehat{e}$ are interior. Since $\widehat{s}$ is interior by assumption, the equilibrium is fully interior in this interval.

The upper extreme of values of q_m such that the equilibrium is fully interior is always q_1^e. Therefore, for $q_m \in [q_1^e, 2]$, $e = 1$, and the equilibrium is characterised as in the analogous case of part (i): $e = 1$, $z = 2 - q_m$, and $s = \frac{\delta+(1-\beta)(2-q_m)}{\delta K}$ (third line of both part (ii) and part (iii)).

The lower extreme of the fully interior equilibrium interval is clearly given by the maximum between $q_1^{\tilde{z}}$ and q_0^e. Using equations (31) and (33), we find that $q_1^{\tilde{z}} > q_0^e$ is equivalent to $\Gamma > -(1-\beta)\delta K$.

In **part *(ii)*,** $\Gamma > -(1-\beta)\delta K$, so that the fully interior equilibrium arises for $q_m \in [q_1^{\tilde{z}}, q_1^e]$ (second line of part (ii)).

For $q_m \in [0, q_1^{\tilde{z}}]$, $z = 1$, and the equilibrium is characterised as in the analogous case of part (i): $z = 1$, $e = q_m$, and $s = \frac{1+\delta-\beta}{\delta K}$ (first line of part (ii))

In **part *(iii)***, $\Gamma < -(1-\beta)\delta K$, so that the fully interior equilibrium arises for $q_m \in [q_0^e, q_1^e]$ (second line of part (iii)).

For $q_m \in [0, q_0^e]$, $e = 0$, so that $z = 1 - q_m + e$ implies $z = 1 - q_m$. The equilibrium level of s must solve condition (A2.4) with $z = 1 - q_m$, yielding: $s = \frac{\delta + (1-\beta)(1-q_m)}{\delta K}$ (first line of part (iii)). ■

Proof of Proposition 3. The essence of the proof is showing that each of the possible pairs (one for the integration and one for the separation regime) of regulation possibility loci has the property that the locus for the separation regime always lies on the left of the locus for the integration regime for every $0 \leq e < 1$. In view of Proposition 2, this implies that the separation regime entails less over-investment in specificity for any value of the effort for quality (except the maximum) the regulator might want to implement by setting the minimum quality standard. Building upon proposition 1 (over-investment in specificity in both regimes), it is sufficient to show that in any parameter region of the model, by moving q_m, the regulator can implement a lower (or, at least, not higher) level of specificity in the separation rather than in the integration regime for any level of effort for quality she would like to achieve.

According to Lemma 1, we can have three alternative shapes for the second best $(s, e) - locus$ of each regime (made explicit in our discussion of proposition 1). Hence, abstracting from the consistency of the parameter conditions, we can in principle combine the integration and the separation loci in nine different ways. We proceed by distinguishing two classes of cases. In the first class, the separation locus has a downward sloping portion (corresponding to fully interior equilibria), as in the two diagrams of Figure 1, while the integration locus can either have a downward sloping portion (like in the two diagrams of Figure 1) or be inverted-L shaped (i.e., the shape arising from the first part of Lemma 1). This class comprises six of the nine possible cases.

We show that, in all cases of this class, the separation locus always lies on the left of the integration locus (except for $e = 1$, where the two loci overlap). Assume first that also the integration locus has a downward sloping portion. Then, using equation (34) with $\delta = 1$ for the separation and $\delta = \beta$ for the integration regime, it is easy to check that the linear downward sloping portion is flatter under integration (recall that $\beta < 1$) and intersects the line $e = 1$ for a higher value of s. Then, irrespective of the shape of the remaining portions of the two loci (i.e., a vertical segment, as in the right diagram of Figure 1, or an horizontal segment lying on the $e = 0 - line$, as in the left diagram of Figure 1), the separation locus is entirely on the left of the integration locus if

$$\frac{1}{\beta K} > \frac{2 - \beta}{K},$$

where the left (right) hand side of the inequality is the measure of the segment $0b$ (resp., $0a$) in the diagrams of Figure 1 (by Lemma 1). This inequality is indeed always satisfied for $\beta < 1$.

There are two other possible cases in the first class: the integration locus is inverted L-shaped, while the separation locus can either be as in the left or as in the right diagram of Figure 1. If so, the inequality above is clearly sufficient for the separation locus always being on the left of the integration locus.

The second class consists of the three cases where the separation locus is always inverted-L shaped, and the integration locus has, alternatively, an inverted L-shape, the shape shown in the left or in the right diagram of Figure 1. If both loci are inverted L-shaped, the inequality above is again sufficient to prove the statement. The remaining two cases could, in principle, invalidate the statement. In these cases, the separation locus is inverted L-shaped, that is, the level of s is fixed at $\frac{2-\beta}{K}$ for any $e < 1$. The integration locus, on the contrary, always exhibits a downward sloping portion (corresponding to fully interior equilibria). We prove, however, that the parameter condition for a fully interior equilibrium in the integration regime is inconsistent with the parameter condition required for the separation locus to be inverted L-shaped, so that these cases are impossible. From parts (i) and (ii) of Lemma 1, a fully interior equilibrium in the integration regime requires:

$$\beta\,(1-\beta)\,[(R-1)K + C] - [\beta K - C\,(1-\beta)^2] < 0.$$

On the other hand, the separation locus is inverted L-shaped only if the condition of part (i) of Lemma 1 holds when $\delta = 1$, that is:

$$(1-\beta)\,[(R-1)K + C] - [K - C\,(1-\beta)^2] \geq 0.$$

Hence, it should be:

$$(1-\beta)\,[(R-1)K + C] - K + \frac{C\,(1-\beta)^2}{\beta} < 0,$$
$$(1-\beta)\,[(R-1)K + C] - K + C\,(1-\beta)^2 > 0,$$

which is impossible since $\frac{C(1-\beta)^2}{\beta} > C\,(1-\beta)^2$ as $\beta < 1$. ■