AN ASSESSMENT OF THE DESIRABILITY OF ON-TRACK COMPETITION: THE IPSWICH-LONDON ROUTE

ANTHONY GLASS[#]

Institute for Transport Studies

University of Leeds

Presented to the First Conference on Railroad Industry Structure, Competition and Investment, Institut D'Economie Industrielle, Toulouse, November 7-8 2003

ABSTRACT

The desirability of competition in the market within the British passenger rail industry is still being fiercely debated. The Office of the Rail Regulator (ORR) and the Strategic Rail Authority (SRA) believe that it is in the public interest to moderate the level of competition in the market. The level of competition in the market is constrained by four barriers to service competition: the mechanism for the moderation of competition; the approach to track access charging; the content of franchise agreements; and capacity constraints. The size of the barriers should reflect the desirability of competition in the market. In this paper, a disaggregate game theoretic model is incorporated into the competition simulation model PRAISE to assess the desirability of more or less competition on the Ipswich-London route. Both alternatives are compared with a monopoly on the route.

Key Words: moderation of competition; simultaneous repeated games; reputation and PRAISE

1. INTRODUCTION

The change in social welfare should be the basis of an assessment of the desirability of more or less competition in the market. More or less competition in the market is taken to be a higher or lower service frequency so the magnitude and possibly the direction of a change in social welfare when there is more or less rivalry between competing operators will depend on the accompanying

[#] Anthony Glass is a fourth year PhD student at the Institute for Transport Studies, University of Leeds under the supervision of Professor Chris Nash and Dr Jeremy Toner.

The PhD research conducted by the author has been funded by an economics Economic and Social Research Council (ESRC) postgraduate research award.

fare changes. Therefore, a game theoretic model is built to forecast the equilibrium in fares. The game theoretic model has several salient features: it operates at the level of a service frequency change as opposed to once and for all entry; two versions of the chain-store game are played simultaneously; it has a leader-follower framework; it permits repeated leadership; and the model is characterised by one-sided imperfect information. The leader is uncertain about the response of the follower because he has imperfect knowledge about the evolution of past play. Imperfect information on behalf of the follower in conjunction with a repeated framework gives rise to a reputation effect for the follower. Imperfect information on behalf of the follower about service frequency changes by the leader is not modelled because operators do not seek to build a reputation for higher service frequencies and an operator certainly does not want a reputation for playing service cuts. Interviews with a number of senior managers in the industry suggested that operators take a short term view of service frequency changes because each change is assessed on its own merits. Higher service frequencies would be a part of a long term strategy if there was reputation building at work.

A round is played in the game theoretic model when there is a service frequency change. There are however a number of barriers to service competition. Once these barriers have been outlined, competition in the market within the British passenger rail industry is briefly analysed to suggest entry strategies that may be able to overcome these barriers. The game theoretic model is then presented in detail. Attention then turns to the empirical analysis beginning with the case study and data requirements. The empirical methodology is then described. It focuses on the approach that is used to solve the two rounds that are played in the empirical analysis. The important results are presented and analysed, and then some conclusions are made.

2. BARRIERS TO SERVICE COMPETITION

Competition on British passenger rail routes is moderated to promote higher franchise bids and private sector investment. It has been noted above that competition in the market is moderated by the mechanism for the moderation of competition; the approach to infrastructure access charging; the content of franchise agreements; and capacity constraints. At present, capacity constraints represent the biggest barrier to service competition followed by the mechanism for the moderation of competition, then the approach to infrastructure access charging and finally, the content of franchise agreements. Each of the four constraints is now discussed in turn.

2.1 Mechanism for the Moderation of Competition

The mechanism for the moderation of competition regulates the level of open access entry. It does not apply to competition between franchises that overlap. In other words, the SRA is free to redraw the franchise map and create more competition between neighbouring franchises. Specifically, the mechanism for the moderation of competition relates to the level of new open access entry. Therefore, the mechanism for the moderation of competition is not a barrier to a higher service frequency on a flow that an open access entrant already serves.

The ORR consulted on the approach to the moderation of competition that should be adopted at privatisation (ORR, 1994a). The responses to the consultation document did not favour a particular approach so the Rail Regulator devised the mechanism for the moderation of competition (ORR, 1994b). The mechanism for the moderation of competition was initially a two stage formula.

Stage I commenced after the final franchise was awarded in March 1997. It constituted almost complete protection from on-track competition. The mechanism for the moderation of competition protects the markets that a Train Operating Company (TOC) serves. A market is identified in terms of a point-to-point flow. In Stage I, a point-to-point flow was defined as a regular scheduled service between two stations. Each direction of travel was classified as a single flow. TOCs nominated material point-to-point flows for exclusivity from on-track competition. Basically, a flow was material in Stage I when the farebox revenue associated with the flow was greater than or equal to 0.2% of the total revenue allocated to the operator on all nominated flows.

Although Stage II was expected to apply from the beginning of April 1999, it effectively applied from the 1999 winter timetable to give the Rail Regulator time to assess the revised list of nominations from each TOC (ORR, 1998). The restriction on the level of on-track competition in Stage II was less prohibitive than Stage I because the materiality threshold increased to 20% of the total farebox revenue allocated to the nominated flows. In Stage II, a franchisee had to make sensible nominations because if a TOC nominated too few flows then a number of high value flows would be completely unprotected. Conversely, if a TOC nominated too many flows, entry on some high value flows would not trigger the materiality threshold. When a franchise consists of a small number of high value flows like the Midland Main Line franchise, only the high value flows should have been nominated because entry on any of the high value flows would trigger the materiality threshold. The highest value flow in the Cross Country franchise apparently yields less than 0.5% of total revenue (Preston et al, 1999). If all the flows in a franchise are relatively low value flows all the flows should have been nominated. The Rail Regulator however, reserved the right to apply Stage II as and when he saw fit to ensure that open access entry did not have a significant impact on the value of franchises (ORR, 1998).

Stage II was initially scheduled to end on 1 April 2002 and be replaced by Stage III. Stage III was to be an incremental change to Stage II (ORR, 1994b). The Rail Regulator however, extended Stage II beyond 31 March 2001 for an unspecified period of time (ORR, 2001a). Stage II will apply until the format of Stage III has been established and a date for its implementation has been set. Shortly after the statement was released to extend Stage II, a format for Stage III was proposed (ORR, 2001b). The format had four important features: (i) protection from on-track competition under Stage II would not apply beyond 31 March 2002; (ii) flows that are protected from on-track competition beyond 31 March 2002 will continue to receive the same level of protection for the agreed period in the track access contract; (iii) nominations for the protection from open access entry in Stage III may be approved to secure investment or prevent cherry picking; and (iv) open access entry on a flow that does not have contractual protection from on-track competition in Stage III would not necessarily be permitted.

The Rail Regulator proposed to publish his draft conclusions on the format of Stage III in autumn 2001 and a policy statement was expected to follow in early 2002. It was anticipated that flows would be nominated for protection in early 2002 and flows would be approved in the summer of that year. The ORR has been unable to keep to the above timetable because the Strategic Plan was published on 14 January 2002 (SRA, 2002a). To avoid a conflict, policy on the moderation of competition must be consistent with the Strategic Plan.

The draft conclusions on the format of Stage III were actually published in the summer of 2003 (ORR, 2003). There were no real changes to the proposed format. The overriding feature of Stage III was the increase in discretionary power that the Rail Regulator would have at his disposal. There would be no materiality threshold in Stage III above which TOCs were protected from on-track competition. Therefore, TOCs would be given no contractual protection from on-track competition in Stage III. The proposed length of Stage III was not specified. It was expected that Stage III would continue until the policy on the moderation of competition needed to be changed. The ORR plans to issue a policy statement by the end of the year on the format and duration of

Stage III. The policy statement will also set the date for the implementation of Stage III. In conclusion, Stage III may represent a bigger barrier to open access entry than Stage II because the Rail Regulator felt that there was no need to give a TOC contractual protection from open access entry, as there would only be a limited number of new cases in the future. This may be because the Rail Regulator plans to use his additional discretionary power to further limit the scope for open access entry. Alternatively, it may be capacity constraints that represent the biggest barrier to open access entry as opposed to Stage III because there are very few paths available on the most profitable routes.

2.2 Infrastructure Access Charges

Within the track access agreement between an operator and the UK infrastructure concern, the charge that is levied for the access rights is specified. TOCs and open access entrants are not charged in the same way. The infrastructure access charge for a TOC is a two-part tariff, consisting of a fixed element and a variable element (ORR, 1994c). The fixed element consists of avoidable costs and an allocation of joint costs, and the variable element is a marginal cost based on wear and tear costs. Moreover, the performance regimes in franchise agreements specify payments and bonuses according to criteria on punctuality and cancellations. Initially, the fixed element was 91% of the total charge. The fixed element fell to 80% of the total charge when 50% of the congestion charge was added to the variable element (ORR, 2000). The variable element however, is still below marginal social cost because besides the remaining 50% of the congestion charge, there is no scarcity charge, environmental charge or external accident charge. The variable element was held below marginal social cost to give TOCs an incentive to increase service frequencies. This represented a smaller barrier to service competition for franchises that overlap than economic theory would suggest.

Initially, the charging regime for open access entrants was a two-part tariff. The variable element has always been equivalent to the corresponding element in the access charge for TOCs. The fixed element was a negotiated contribution to the fixed costs of the incumbent TOC and was based on the ability to pay (DTp, 1993). The open access charge now only consists of a variable element (ORR, 2000). If the open access charge were based on the efficient component pricing principle (Baumol, 1983) the charge would represent a considerably bigger barrier to entry. The efficient component pricing principle contains three elements long run marginal social cost, a contribution to the fixed costs of the incumbent and compensation to the incumbent for loss of business. The open access charge only corresponds to the first element of the efficient component pricing principle and this is even below long run marginal social cost. If the recent draft conclusions on Stage III become policy, open access entry would be permitted when the incumbent is protected from on-track competition if the entrant compensated the incumbent for loss of business.

2.3 Franchise Agreements

There are a number of aspects of a franchise agreement that may constitute a barrier to service competition: the Passenger Service Requirement (PSR); fares regulation; support payments; inter-operator arrangements; and the formation of coalitions.

The PSR of a TOC is on a route-by-route basis and specifies the following: frequency of trains; stations to be served; maximum journey times; first and last trains; weekend services; through services; bus links; and load factors and peak train capacity. The PSR obligations of a TOC can prevent entry on a new route or the introduction of additional services on an existing route if a TOC wants to use rolling stock that is being used to fulfil PSR obligations. Jones (2000) described the PSR as a total barrier to exit. This is not strictly true because operators can trade PSR obligations. In practice, TOCs rarely do so.

The franchise agreement regulates fares. As a result of the review of fares regulation, the price cap on regulated fares will be relaxed (RPI+1% as opposed to RPI-1%) for three years from the January 2004 fares change and all fares will be regulated by the basket approach (SRA, 2003a). As of the January 2004 fare change, every TOC on the network will have an additional protected fares basket. Protected fares were previously regulated individually. The increase in an individual fare in a basket can exceed the price cap on the basket if this is compensated for by changes in at least one other fare change. To prevent excessive increases in individual fares, a price cap is placed on each fare within a basket. The price cap on an individual fare within a basket was also relaxed; RPI+6% will replace RPI+2%. Although the new policy on fares regulation gives TOCs more flexibility with regard to the size of fares increases it still constitutes a barrier to service competition because it restricts competitive freedom. This is particularly the case when markets are price inelastic.

The subsidy from or the payment to the SRA is described in detail in a franchise agreement. It is imperative that the SRA gives and receives the appropriate subsidies and payments. An excessive subsidy or an insufficient payment will constitute excessive regulation of competition in the market because it will put a franchise in a position to cross subsidise a route to deter a higher service frequency or invoke a lower service frequency. In theory, a TOC can renegotiate a higher subsidy or lower payment with the SRA midway through a franchise term. To the best of my knowledge this has never been successfully done.

The franchise agreement represents a commitment from a TOC to co-operate and maintain a number of benefits across the passenger rail network. One of these network benefits is inter-available ticketing. The lead operator on flow (usually the operator with the biggest commercial interest on a flow) is responsible for setting inter-available fares. The lead operator does not have to consult competing operators about inter-available fares. The lead operator can therefore set inter-available fares to maximise his own utility. The SRA are willing to drop the inter-available ticketing requirement if more intense price competition and greater service diversity would be beneficial. The inter-available ticketing requirement on the Gatwick Airport-London flow and some of the intermediate flows has been lifted. At present every other flow on the network has a lead operator. It is standard practice for an entrant to sell operator specific walk-up tickets. Entry is often attractive because cheap operator specific tickets can be used to compete with expensive inter-available tickets. If the SRA however, began to lift the inter-available ticketing requirement on more routes this may deter entry.

More than one company can jointly own a franchise. For example, Stagecoach bought 49% of the West Coast franchise that Virgin won at privatisation. Joint ownership of a franchise, especially by two large companies like Virgin and Stagecoach, may be a barrier to service competition because joint ownership will give a TOC access to more finance, which would probably limit the ability of an entrant to compete.

2.4 Capacity Constraints

The infrastructure authority is responsible for a lot of the investment on the British rail network. When resources are scarce the infrastructure authority like any other organisation will prioritise. As a result, there will be situations where capacity constraints prevent more competition in the market because work on the infrastructure has not been completed, been postponed or even cancelled. Many commentators have regularly noted the lack of investment by the infrastructure authority. This was very much the case when Railtrack was the infrastructure concern. The shortfall in investment has been partially addressed by Special Purpose Vehicles (SPVs). SPVs are joint investment projects between the infrastructure authority and one or more operators. At the beginning of March 2001 Virgin Trains and Railtrack entered into an agreement to carry out £200m of work on the Cross Country Line. The infrastructure authority still owns the assets that are developed by SPVs but the size of the contribution by an operator will depend on the return that it can make on the investment. To a large extent this governed by the length of the franchise. In the case of Virgin Cross Country, its contribution to the above SPV was substantial because the current Cross Country franchise only expires in January 2012.

The Strategic Plan set out to tackle the shortage of capacity on the network. Specifically, the SRA proposed to make more efficient use of the existing capacity and promote private sector investment in the railways. A series of franchise mergers were proposed to make more efficient use of existing capacity. As a result of the mergers, there will only be one TOC running to each London terminal. When more than one operator runs to a London terminal there will invariably be a mix of rolling stock running on the same route. Capacity may not be allocated efficiently because the slower stock may slow down the faster stock. It was however noted in the Strategic Plan that the SRA was not just seeking to make more efficient use of existing capacity by merging franchises. The SRA outlined plans to take a more active role in the allocation of capacity by becoming more involved in timetable planning (SRA, 2003b). According to Nash (2003) the new approach to the allocation of capacity would be more efficient because timetables would be more flexible. At present, timetables are specified in the PSRs. Reallocating capacity between TOCs is a difficult and time consuming task because it involves renegotiating franchise agreements and track access agreements with at least two TOCs. Under the new system a timetable will be mutually agreed between the SRA and the operator. Efficiency gains in the short term will facilitate long term gains. Efficient allocation of existing capacity in the short term will ensure that private sector investment gravitates to the parts of the network where it will be of most benefit.

3. ENTRY STRATEGIES

There are a number of possible entry strategies that could be used on a British passenger rail route. In research that is ongoing, it is argued that an operator will use the same strategy across the whole of its business to fully exploit economies of traffic density. This may or may not mean that an operator uses the same entry strategy on more than one route; an entry strategy of course depends on the strategy of the incumbent. Preston et al (1999) suggest a number of entry strategies e.g. head-on competition, cherry picking, niche market entry etc. Head-on competition in the context of this research is taken to be a regular service throughout the day using the same route and similar rolling stock as the incumbent. Under the present policy on competition in the market the ORR would almost certainly take the stance that the public interest is not best served by head-on competition between an open access entrant and a franchisee. This view is based on the record of open access entry in the industry to date. Although there has only been one case of open access entry, the rivalry between Hull Trains and Great North Eastern Railway (GNER) could hardly be described as head-on competition. Moreover, it has already been noted above that the SRA want to merge franchises and limit the number of cases of where there is competition between franchises. Specifically, the SRA want to limit head-on competition. The SRA are however keen to promote competition at the margins (SRA, 2003b).

It is therefore argues that an open access entrant or a TOC that wants to extend the boundary of his franchise should adopt niche market entry, vertical product differentiation or horizontal product differentiation to satisfy the public interest criteria that the ORR and the SRA use to assess the desirability of competition in the market. Niche market entry is intended to meet demand from marginal customers that is not met by the incumbent. The incumbent may not meet the demand for several reasons; it may be non-core business, pose management difficulties and/or be insufficiently profitable. Charter trains luxury services such as the Orient Express, night services, motor-rail services and direct services between towns and cities that previously did not have such services in the recent past are all forms of niche market entry. The direct Hull-London service that Hull Trains operates is an example of the latter.

Vertical product differentiation in the context of the railways refers to service quality competition. Horizontal product differentiation on the other hand is at work in a passenger rail industry when there is a spatial element to competition. A textbook definition of horizontal product differentiation would encapsulate temporal competition but there is necessarily temporal differentiation between services on the railways for safety reasons. Horizontal product differentiation is therefore taken to be competition by route. There is competition via vertical and horizontal product differentiation on the Birmingham-London flow. The incumbent on this flow is Virgin West Coast. Silverlink competes with Virgin using vertical product differentiation because both operators serve Birmingham New Street and London Euston with different rolling stock. Virgin operates with Class 390 electric tilting trains, which have a maximum speed of 140 mph. The Virgin services however only run at 110 mph because of speed restrictions on the line. Silverlink on the other hand use 100 mph Class 321 electric multiple units. Moreover, Chiltern uses horizontal product differentiation to compete with Virgin because Chiltern serves Birmingham Moor Street and London Marylebone.

Unfortunately, the classification of cases of competition in the market according to entry strategies is not so simple because a number of entrants play mixed strategies. When an entrant is using mixed strategies the case is classified according to the most dominant strategy. For example, the competition between Chiltern and Virgin West Coast on the Birmingham-London flow is classified as horizontal product differentiation but Chiltern and Virgin use different rolling stock so there is also an element of vertical product differentiation in this case.

4. GAME THEORY

The theoretical model is presented in several stages. In the first stage, the solution concept that is used is defined intuitively. In the second stage the assumptions of the model are outlined and in the third stage the two player games that are presented are solved. Specifically, a two player lower service frequency game and a two player higher service frequency game are presented and solved. In research that is in the final stages of completion a second version of the lower service frequency game has been built. Furthermore, in this work in progress, each two player game is extended to rivalry when there are multiple entrants. Returning to the game theory that is presented in this paper, it is assumed that the two version of the chain-store game are played simultaneously. A round of a repeated game is played when there is a timetable adjustment in terms of a service frequency change. A service frequency change has ‘good’ potential for a player in period *t* when (1) applies:

$$\exp(\Delta \pi_t) > 0 \tag{1}$$

It has been noted above that the repeated games in the theoretical model are characterised by one-sided imperfect information because the leader cannot make an exact assessment of past play. There are two sources of imperfect information: (i) a non-vanishing element of doubt that a follower will behave as he has done in the past if an identical situation arose and (ii) the a commercially sensitive information that only the follower has access to.

There will be doubt in the mind of the leader that the follower will behave as he has done in past play in identical circumstances because managers get replaced. Very often a successor has a different managerial philosophy, which gives rise to inconsistent responses over time. Furthermore, the leader will not have access to commercially sensitive information. For example, a prospective leader will not know for certain if a fare cut is a competitive response to a service frequency change or part of general company policy. Cheap APEX tickets is part of general company policy at Virgin to attract passengers from other modes. It is not a competitive response to Chiltern and Silverlink service frequency changes between Birmingham and London.

4.1 Sequential Equilibrium

The concept of sequential equilibrium is the Nash equilibrium refinement that is used because the game theoretic model draws upon work by Kreps and Wilson (1982a). Intuitively, an outcome is a sequential equilibrium if each player has selected a strategy that is optimal for the remainder of the game based on a rational interpretation of previous play and a rational expectation of future play. Kreps and Wilson (1982a) note that an outcome is a sequential equilibrium if the following criteria are fulfilled:

(i) a player attaches a probability assessment to the nodes in his information set according to a player's subjective view of outcomes in previous rounds;

(ii) the probability assessments are computed using Bayes theorem when it applies and are in line with the hypothesised equilibrium strategies; and

(iii) from every information set in the game, each player's behaviour is consistent with sequential rationality. Therefore, a strategy can only be optimal for the remainder of the game if it is based on a rational perception of future behaviour.

The reader is directed to Kreps and Wilson (1982b) for a formal exposition of a sequential equilibrium. Kreps and Wilson prove that all sequential equilibria are Nash.

4.2 The Repeated Game Across Subsets

The first version of the chain-store game is the repeated game across subsets. This is the most common version of the chain-store game because there is entry in sequence in a different subset in the context of this model. Suppose that Q is the finite set of markets that the incumbent operates in. The incumbent operator is the lead operator on a flow. The lead operator on a flow is the operator with the biggest commercial interest. $|Q|$ will vary according to the level of disaggregation that the game theoretic model operates at. Suppose the incumbent operates on W flows where a flow is an origin-destination pair in one direction. Let there be K ticket types on each flow, where $K \varepsilon [2, \infty)$. Therefore, $|Q|=WK$. If the analyst elected to operate at the most aggregate level of average fare, there will be at least an average single fare and an average return fare for travel on a flow. Season tickets are not available on all flows.

Assume that Q consists of P distinct subsets where there is initial entry in sequence, so $P \varepsilon [2, \infty)$. $W \varepsilon [4, \infty)$ must be the case because the incumbent will run in both directions of travel. The Gatwick Express is the incumbent on two flows, Gatwick Airport-London and London-Gatwick Airport. Therefore, W=2 for the Gatwick Express, so the Gatwick Express is the only TOC that the repeated game across subsets does not apply to.

Let p denotes any of the P subsets where $|p| \varepsilon [4, \infty)$ because each p will contain at least an origin-destination pair in each direction and even if the model is used at the most aggregate level there will be an average single fare and an average return fare in each direction. It is assumed that entry occurs in each p in sequence. Entry in the repeated game across subsets corresponds to the initial level of entry in each p. The initial level of entry refers to the initial timetable of the first entrant to each p.

There are M+1 players in the repeated game across subsets. There are M entrants and an incumbent. Let m denote any of the M entrants where $M \varepsilon [1, P]$. If $M<P$ at least one entrant will have re-entered. The P subsets indexed from 1,…,P. Similarly, the M entrants are indexed from 1,…M according to the sequence of entry. m_1 enters p_1 but since re-entry is possible nothing can be said about the identity of the entrant in any other subsets if in fact there is more than one entrant in the repeated game across subsets.

4.3 The Repeated Game Within a Subset

The second version of the chain-store game is not cited as regularly. The second version in the context of the model that is presented here is a repeated game within each subset. Each round of the repeated game across subsets is the first round of the repeated game within a subset.

Suppose that V denotes the number of entrants in the repeated game within each subset, where $V \varepsilon [1, \infty)$. There are V+1 players in total in a repeated game within a subset. The entrants in each repeated game within a subset are indexed from 1,…,V. Let v denote any of the V entrants. v_1 corresponds to an m in the repeated game across subsets but v_1 will not correspond to a distinct m if there repeated entry in the repeated game across subsets. If V=1 however, the rounds in the repeated game within subset p will be two player rounds.

In each p, it is assumed that there are R_p^c rounds that the incumbent can initiate where c is used to denote the incumbent. $R_p^c \varepsilon [2, \infty)$ because it is assumed that the incumbent initiates at least two rounds. Therefore, at least entrant 1 in the repeated game within subset p will follow at least twice. Similarly, let there be R_p^v possible rounds in each p that entrant v can initiate where $R_p^v \varepsilon [2, \infty)$. It is therefore assumed that each entrant in the repeated game within a subset will initiate at least two rounds. Even though the repeated game across subsets does not apply to the Gatwick Express, it is assumed that the repeated game within a subset is played on the Gatwick Airport-London route because $R_p^c \varepsilon [2, \infty)$ and $R_p^v \varepsilon [2, \infty)$.

Let r_p^c and r_p^v denote any round in R_p^c and R_p^v. Moreover, let r_p^c* be the last round that the incumbent initiates. Also, let r_p^v* be the last round that entrant v initiates. Therefore, r_p^c* is the last service frequency change that has 'good' potential for the incumbent and r_p^v* is the last change that has 'good' potential for entrant v. r_p^c* and r_p^v* are off course dependent on the sequence of play. The number of rounds in the repeated game within subset p is given by the following expression:

$$S = r_p^c * + \sum_{v=1}^{V} r_p^v * \quad (2)$$

It is assumed that there is game playing in the *P* subsets over a finite number of periods *T* where $T \varepsilon [5, \infty)$. $T \geq 5$ must be the case because $P \geq 2$, $r_p^c * \geq 2$ and $r_p^v * \geq 2$. The general expression for *T* is as follows:

$$T = \max[S + (p-1)] \quad (3)$$

4.4 Reputation Acquisition

Let *t* denote any of the *T* periods. In every *t* there will be at least one service frequency change. Before each service frequency change the leader does not know how the follower will respond. Therefore, the leader does not know if he should move. To establish if he should move it is assumed that the leader assesses the reputation of the follower using past play. Obviously, if a follower under assessment has never responded before he will not have a reputation. *Ad hoc* assumptions are now made about the acquisition of a reputation by a follower. These assumptions depend on the level of disaggregation of the analysis.

If the model is used at the most aggregate level of average fare, it is assumed that the leader makes two or three assessments of the follower's reputation according to changes in the average single fare, average return fare and possibly the average seasonal fare if there is a season ticket market on a flow where the leader is planning the service frequency change. If there are *J* types of assessment where *j* is used to denote any of the *J* assessments, when the model is used at the most aggregate level J=2 or J=3. Alternatively, if the analyst uses the model at the most disaggregate level it is assumed that the number of assessments of the follower's reputation will be equal to the number of like for like ticket types that the leader and the follower will compete with if the service frequency change is implemented.

Each assessment is based on the response of a follower to a service frequency change. A leader will consider responses to higher service frequencies and lower service frequencies separately because the response of the follower in past play may vary according to the direction of a service frequency change. For simplicity, a distinction is not made between a lower and higher service frequency change in the description of reputation acquisition that follows. Since the leader does not know how the follower will respond so the likelihood that the follower will play fight or accommodate in period *t* in the *jth* type of market is assessed with probabilities δ_{t_j} and $(1-\delta_{t_j})$ respectively. The follower under assessment is assumed to have played fight in past play if he responded to a service frequency change with a fare cut. Conversely, the follower is assumed to have played accommodate in past play if the fare change in response to a service frequency change was not a fare cut. δ_{t_j} is estimated using Bayes theorem and depends on the overall fare elasticity for a flow where the leader is planning a service frequency change. A prospective leader would know the overall fare elasticity of demand from an industry model like MOIRA. The fare elasticity for a flow is part of an assessment of a follower's reputation because to some extent it governs how aggressive a fare cut in past play actually was. The leader has imperfect information because he does not have access to the follower's *jth* type elasticity of demand.

A leader is assumed to use equation (4) to estimate δ_{t_j}, when the overall fare elasticity of demand on a flow where there the leader is planning the service frequency change is elastic:

$$\delta_{t_j} = \Pr(Fare\ Cut \mid Elastic\ Flow) \tag{4}$$

$$\delta_{t_j} = \frac{\Pr(Fare\ Cut \cap Elastic\ Flow)}{\Pr(Elastic\ Flow)}$$

$$\delta_{t_j} = \frac{\Pr(Elastic\ Flow \mid Fare\ Cut) * \Pr(Fare\ Cut)}{\Pr(Elastic\ Flow \mid Fare\ Cut) * \Pr(Fare\ Cut) + \Pr(Elastic\ Flow \mid No\ Fare\ Cut) * \Pr(No\ Fare\ Cut)}$$

A similar expression applies when the overall fare elasticity of demand on the flow of interest is inelastic.

Pr(*Elastic Flow*|*Fare Cut*) and Pr(*Fare Cut*) are weighted probabilities so Pr(*Elastic Flow*|*No Fare Cut*) and Pr(*No Fare Cut*) will be weighted probabilities as well. However, attention now turn to the computation of Pr(*Elastic Flow*|*Fare Cut*) and Pr(*Fare Cut*).

Suppose that there have been a finite number of responses of the *jth* type in past play, G_j. Now suppose that the leader assesses H_j battles where $H_j \leq G_j$. It is assumed that if $H_j < G_j$ the H_j battles are randomly selected. If $H_j < G_j$ the information set of the leader is even more imperfect because he has less information about the evolution of past play than he might have. When h_j is used to denote any of the H_j battles, then Pr(*Fare Cut*) can be calculated as follows. Pr(*No Fare Cut*)=1-Pr(*Fare Cut*):

$$\Pr(Fare\ Cut) = \frac{1}{H_j} \sum_{h_j}^{H_j} \psi_{h_j} \eta_{h_j} \tag{5}$$

where:

> ψ_{h_j} *is a* (0,1) *dummy variable that takes a value of* 1 *if in competitive battle* h_j *the follower under assessment played fare cut and*
>
> η_{h_j} *is a weight that measures the aggressiveness of a fare cut.*

η_{h_j} is calculated with the help of a reference fare cut, the biggest *jth* type fare cut in past play. Therefore, $\eta_{h_j} = \theta_{h_j} / \theta^*$, where θ_{h_j} is the fare cut of the *jth* type in the battle h_j as a percentage and θ^* is the biggest *jth* type fare cut in past play as a percentage by the player under assessment.

The aggressiveness of a fare cut also depends on the overall fare elasticity for the flow that battle h_j corresponds to. Now suppose that the follower under assessment played fight in F_j battles where, $F_j \le H_j$. If f_j is used to denote any of the F_j battles then Pr(*Elastic Flow*|*Fare Cut*) is calculated using (6). Pr(*Elastic Flow*|*No Fare Cut*) can be calculated in a similar way:

$$\Pr(Elastic\ Flow \mid Fight) = \frac{1}{F_j} \sum_{f_j}^{F_j} \tau_{h_j} \rho_{f_j} \tag{6}$$

where:

> τ_{h_j} *is a* (0,1) *dummy variable that takes a value of* 1 *if the flow that corresponds to competitive battle* f_j *is elastic and*
>
> ρ_{f_j} *is a weight attached to the elasticity of the flow that corresponds to battle* f_j.

ρ_{f_j} is calculated with the help of a reference overall fare elasticity. The overall fare elasticity for the least elastic flow on which there has been fare cut in response to a service frequency change in past play is taken to be the reference elasticity. Therefore, $\rho_{f_j} = \theta^* / \theta_{f_j}$, where, θ_{f_j} is the overall absolute fare elasticity of the flow that corresponds to competitive battle f_j and σ^* is the overall absolute fare elasticity of the reference flow. If however, overall demand on the flow where the prospective leader is considering a service frequency change is fare inelastic ρ_{f_j} will be the inverse of the above expression. In this case, the reference elasticity will be the most inelastic flow where there has been a fare cut in response to a service frequency change in past play.

Since the leader must play higher or lower service frequency in all *J* markets or not move he will calculate an average δ_t across the *J* markets in which he will compete with the follower. The leader will weight δ_t according to Operational Research Computer Allocation of Tickets to Services (ORCATS) shares if the service frequency change is not the initial entry timetable. If the service frequency change is the initial entry timetable δ_t will be an unweighted average because a prospective entrant will not have access to ORCATS ticket shares. Where there is on-track competition on a flow an operator knows the total ORCATS market share of each ticket and his own ticket shares.

The *ad hoc* assumptions that have been made to show how δ_t may be estimated using Bayes theorem are only suggestions based on the information available to the leader. Bayesian updating of the reputation of a follower could just as easily be based on another set of *ad hoc* assumptions.

4.5 Two Player Service Frequency Games

Figure 1*: The Two Player Service Frequency Games in Extensive Form*

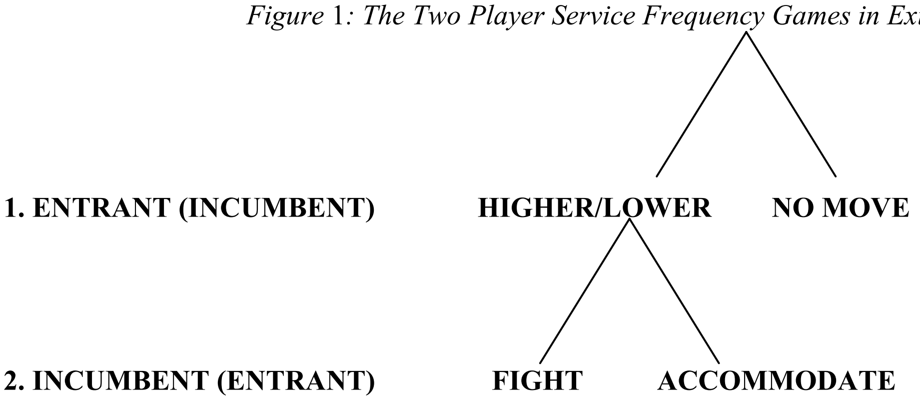


Table 1: Pay-offs for the Two Player Higher Service Frequency Game

(ENTRANT/INCUMBENT, INCUMBENT/ ENTRANT)	**PAY-OFFS**
(No Move, No Move)	(0, 0)
(Higher, Fight)	(Ω, -α+β)
(Higher, Accommodate)	(Ω, -γ)

Table 2: Pay-offs for the Two Player Lower Frequency Game

(ENTRANT/INCUMBENT, INCUMBENT/ ENTRANT)	**PAY-OFFS**
(No Move, No Move)	(0, 0)
(Lower, Fight)	(Ω, α+β)
(Lower, Accommodate)	(Ω, γ)

4.6 Solving the Two Player Service Frequency Games

There are two two player service frequency games, the higher service frequency game and the lower service frequency game. Each game is solved at the level of a market. If the model is used at the most aggregate level, the relevant two player game will be solved at the level of average single fare and average return fare on each flow where there is a service frequency change. At the most disaggregate level, a game will be solved at the level of individual like for like tickets that the competing operators sell. To solve each game it is assumed that the cross fare elasticities of demand between markets are minimal[1].

Either game begins when the potential of the service frequency change that the leader is considering is 'good'. Part of the assessment of the potential of a service frequency change is based on the anticipated response from the follower. It has been noted above that the leader assesses the reputation of the follower to attach probabilities to the two possible responses fight and accommodate. A service frequency change can only have 'good' potential if the follower does not have a sufficiently big reputation for playing fight. If the follower has a big enough reputation to deter a service frequency change the leader will not move. Formally, the reputation

[1] Strong assumptions are frequently required to allow the modelling to proceed. Baumol and Bradford (1970) made the same assumption in their classic study of second best pricing.

of the follower for playing fight will not deter the service frequency change if $\delta_t < \delta_t^*$. δ_t^* is a threshold that the leader sets.

It may seem paradoxical that the reputation of a follower can deter the leader from withdrawing services from his timetable. Even if the services that the leader is considering withdrawing are unprofitable, initiating the lower frequency game may not be rational if there is high probability that the follower will respond with fight.

The pay-offs for each game are in terms of change in profit for each player. The pay-off to the leader is in terms of Ω in the present period t*. If the follower plays accommodate his pay-off is in terms of γ in period t*. If however, the follower plays fight his pay-off consists of two elements α and β. α is realised in the present period but β is realised over the course of a number of periods beyond the period t*.

The follower invests in his reputation when he plays fight. In theory, as with any other investment, if the follower invests in his reputation it can yield a return or a loss. In practice, a reputation for playing fight will be an asset so it is assumed that β>0:

$$\beta = \begin{cases} \sum_{t=t^*+1}^{t=T} \lambda^{t-t^*} \xi_t \text{ if } T - t^* \geq 1 \\ 0 \text{ otherwise} \end{cases} \tag{7}$$

where:

ξ_t is the estimate of the per period profit retention or profit forfeiture and

λ is the per period discount factor.

In other words, β can only be realised if there is game playing in future periods. There are five relations between α, γ and β. The relationship that applies determines the sequential equilibrium:

(i) $\alpha>\gamma$ and $\beta>0$;

(ii) $\alpha<\gamma$ and $\beta>0$ such that $\alpha+\beta>\gamma$;

(iii) $\alpha<\gamma$ and $\beta>0$ such that $\alpha+\beta<\gamma$;

(iv) $\alpha=\gamma$ and $\beta>0$; and

(v) $\alpha<\gamma$ and $\beta>0$ such that $\alpha+\beta=\gamma$.

Intuitively, the relation between α, γ and β depends on six factors. The follower only knows (a) with certainty. He will have to form an expectation of (b), (c), (d), (e) and (f):

(a) How much does a marginal investment in period t^* cost?

(b) How often will the marginal investment deter higher or lower service frequency?

(c) What will constitute higher or lower service frequency? For example, in the case of higher service frequency, will the deterred frequency be one additional service per day or an hourly service?

(d) Where will higher or lower service frequency be deterred?

(e) When will higher or lower service frequency be deterred?

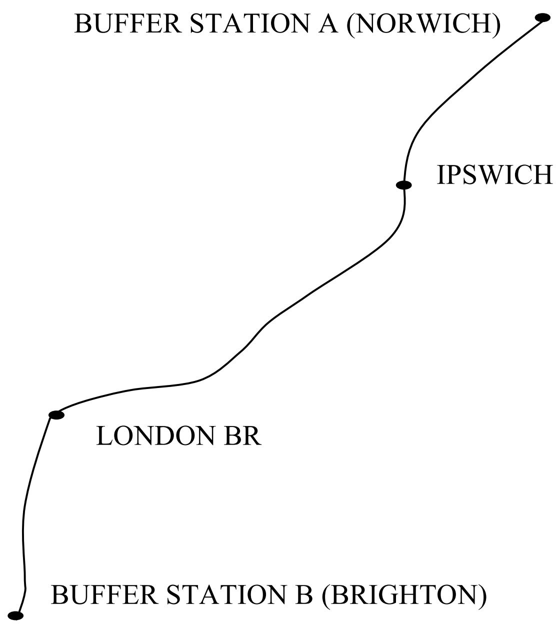
(f) Who will be deterred from playing higher or lower service frequency?

It is assumed that the follower has a long term profit objective which dominates his short term profit objective. The long term profit objective is simply profit maximisation over the course of the T periods. The short term objective is profit maximisation in each round. The short and long term objective may be mutually exclusive in some rounds. When (ii) applies the short and long term objectives are mutually exclusive but the follower plays fight to maximise long term profit if. When however, (i) or (iii) is relevant the short and long term objectives are complementary and the follower plays fight and accommodate. The follower can make a costless contribution to his reputation if (i) holds but if (ii) applies the contribution costs γ-α.

5. CASE STUDY AND DATA

The simplified Ipswich-London BR network in figure 2 is the corridor that is modelled in the empirical analysis to assess the desirability of competition in the market between Anglia Railways and First Great Eastern on the Ipswich and London route. Competition on the Ipswich-London route is a particularly interesting issue at the moment because there will be a monopoly operator Greater Anglia between Ipswich and London as of 1 April 2004. The Greater Anglia franchise will be the amalgamation of Anglia Railways, Great Eastern and West Anglia services from West Anglia Great Northern franchise[2]. Anglia is the incumbent on the Ipswich-London flow because Anglia is the lead operator. The lead operator sets the inter-available fares. The entrant, First Great Eastern competes with a set of operator specific fares:

Figure 2: *The Simplified Ipswich-London BR Network*



In the weekday 2000-2001 winter timetable Anglia and First Great Eastern competed on a number of intermediate flows on the Ipswich-London route (e.g. Colchester-London; Manningtree-London etc). The intermediate flows are ignored for two reasons. First, limiting the number of stations on the corridor minimises the run time of the PRAISE program and second, the stations on the above corridor are restricted to stations that are contained in MOIRA. Anglia and First Great Eastern run to London Liverpool Street but MOIRA only contains a London BR station. The London BR station in MOIRA corresponds to fourteen London stations.

---

[2] First Group submitted a bid for the new Greater Anglia franchise but it was not one of three short listed companies (SRA, 2003c). The three companies were Arriva Trains, National Express and the parent company of Anglia Railways, GB Rail. The shortlisting criteria was the subject of the resulting legal action by First Group against the SRA. First Group dropped the legal action when a bid for GB Rail was accepted and the SRA sanctioned the takeover (Osborne, 2003). The takeover of GB Rail by First Group is discussed in greater detail in research that is ongoing.

Competition between Brighton and London BR is not modelled and there is no competition between Norwich and London. Nevertheless, Norwich and Brighton form part of the network that is modelled because they are the representative stations for the buffer stations. Buffer stations are notional stations that are introduced to the network to estimate demand that feeds into the Ipswich-London flow. Demand is aggregated across the stations in the buffer. There are only four stations in buffer A (Cromer; Norwich; Great Yarmouth; and Lowestoft) because there are very few major MOIRA stations beyond Ipswich. In contrast, there are a large number of stations beyond London that passengers could be travelling to or from. These stations are included in buffer B. There are 159 stations in buffer B. A station within a buffer that is an average distance away from the Ipswich-London BR corridor must act as the representative station. The service level between the representative station and the Ipswich-London corridor must be reasonably frequent. The Anglia Railways fare structure, ticket shares and timetable to and from Norwich is used as a proxy for travel to and from the stations in buffer A. Similarly, the South Central fares structure and timetable together with the total market shares of tickets[3] to and from Brighton are used to represent stations in buffer B.

PRAISE consists of three models (demand, cost and evaluation models). The demand model is a hierarchical logit model that estimates the modal share for rail, splits the rail market up according to ticket type and allocates demand for each ticket type to individual services. The cost model is far simpler. For each operator, a fixed cost, a variable cost per train kilometre and a variable cost per train hour are allocated to the route that is modelled. Finally, the evaluation model estimates the unweighted change in social welfare in terms of the change in total operator profit and the change in consumer surplus when any combination of the following change: journey time; service frequency; number of tickets; restrictions on tickets; stopping patterns and fares. The reader is directed to Johnson and Whelan (2003) for a detailed formal presentation of the PRAISE model.

To make the demand model operational three types of input are needed: network information; modal shares; and information on tastes and preferences. Network information refers to a timetable, fares, ticket shares by operator and base demand for each origin-destination pair. In December 2002, the Institute for Transport Studies, University of Leeds completed a project for the SRA (Whelan, 2002). Under the terms of this project the SRA provided the network

[3] The reference ticket shares for travel to and from Brighton are total market ticket shares because there is competition in the market on the Brighton-London flow between South Central and Thameslink.

information for the above Ipswich-London BR network. The information is used again to carry out a more comprehensive study of competition on the Ipswich-London BR corridor because the work for the SRA only illustrated how PRAISE could be used to simulate competition on a British passenger rail corridor. Written permission from Anglia Railways and First Great Eastern to use the above information in this research has been received.

The SRA supplied: (i) an electronic Wednesday 2000-2001 winter timetable, which applies to the period 24 September 2000-19 May 2001; (ii) 2000-2001 ticket shares by operator; (iii) 1999-2000 annual demand data and (iv) 2000-2001 generalised journey time elasticities of demand for each origin-destination pair. The electronic timetable is used to determine the realistic journey opportunities for travel between each origin-destination pair on the above network. In the empirical analysis, a journey opportunity is assumed to be unrealistic if: (i) there is more than one interchange involved; (ii) the interchange period is less than 7 minutes or greater than 60 and/or (iii) the overall travel time is greater than twice the quickest journey opportunity.

The fares data was taken from the fares manuals for the period 7 January 2001-19 May 2001. The practical modelling focuses on four ticket types, first class return tickets, standard class return tickets, reduced fare return tickets and season tickets, with and without inter-availability as and when it is appropriate. The representative season ticket in the empirical analysis is the 7 Day Season ticket. A 7 Day season ticket is converted into a daily return fare using the typical number of journeys suggested in the CAPRI (computer analysis of passenger receipts) manual. The ticket shares were derived from 2001 ORCATS information. The ticket share information is however inaccurate because the demand data and ticket shares do not include the Anglia Railways *Commuter Club* market[4]. It is safe to say that this market is quite a large so the Anglia season ticket shares will be underestimates.

The annual demand data is in terms of the number of return trips between each origin-destination pair on the network that is modelled. The annual demand data is taken from CAPRI. CAPRI is a set of computer programs that sends information each night over the telephone line about the tickets that have been sold throughout the day (Fowkes and Nash, 1981). Daily shares for a

[4] The *Commuter Club* ticket is valid for travel on all Anglia services unlike a point-to-point season ticket. The ticket is valid for a month and payment is via a monthly direct debit. Compensation payments for delays over 30 minutes on nominated services are paid directly into a passenger's bank account.

typical week were applied to the annual data to obtain demand for a typical Wednesday in 1999-2000.

The 1999-2000 demand data is used to derive demand data for a typical Wednesday in 2000-2001. The 1999-2000 data for each origin destination pair is scaled up or down by the percentage change in passenger journeys for the period 2000-2001. The percentage change in passenger journeys for Anglia Railways and First Great Eastern for the period 2000-2001 is 5.1% and -0.5%, respectively (TAS, 2002). An average percentage change in passenger journeys of 2.7% for the period 2000-2001 is taken from the SRA National Rail Trends. Demand between origin-destination pairs when the origin station is Brighton is scaled up by 2.7%. When the origin station is Ipswich or Norwich demand is scaled up by 5.1% because the percentage change in Anglia passenger journeys is more reflective of total demand for the flows on the Norwich-London corridor.

To establish by how much the rail modal share can grow or contract for each flow, the upper nest of the hierarchical logit model requires a base set of modal shares. The base modal shares for each flow are derived from the shares published by Department of Environment, Transport and the Regions (2000), DETR from hereon. The DETR shares were adjusted to capture the possibility of non-travel and pure generation effects when the rail market grows or contracts. There is however a discrepancy between the modal share information and other data that are needed to make the demand model in PRAISE operational. The modal share information is for the period 1997-1999 but this is not a great cause for concern because the information for January 2001 will be very similar.

For each flow on the network, the tastes and preferences of the individuals in the sample are captured by generalised journey time elasticities of demand and rail service valuations. Generalised journey time elasticities of demand for three ticket types (full fare, reduced fare and seasonal) are used. These elasticities are taken from MOIRA. In the practical modelling it is assumed that the full fare generalised journey time elasticity also applies to first class demand.

The service attribute valuations are based on the valuations in the fourth edition of the Passenger Demand Forecasting Handbook (ATOC, 2003), which from hereon is referred to as the PDFH. Monetary valuations are attached to in-vehicle time, adjustment time, interchange time, the inconvenience associated with changing trains and the discomfort that crowded conditions cause. The monetary valuations in the PDFH are at 1 January 2001 prices and incomes.

The valuation of in-vehicle time is the valuation of the cost to an individual of a minute spent travelling by rail. The value of in-vehicle time varies according to journey purpose, class of travel and distance. An unweighted average is taken across journey purpose and class of travel, so the value of in-vehicle time only varies by journey distance. Adjustment time is simply the difference between a passenger's preferred departure time and his actual departure time and the value of interchange time is based on the value of in-vehicle time according to the method proposed in the PDFH. Accordingly, interchange time is valued at twice the value of in-vehicle time. A related service attribute is the interchange penalty. The interchange penalty is an estimation of the inconvenience associated with changing trains in terms of minutes, which is then converted into a monetary valuation using the value of in-vehicle time.

The final service attribute is an overcrowding penalty. The overcrowding penalty is a monetary estimation of the dissatisfaction that a passenger attaches to standing or sitting on a crowded train. In line with recommendations in the PDFH, the overcrowding penalty is in terms of the value of in-vehicle time. When a passenger is forced to stand or sit on a crowded train, his value of in-vehicle time is inflated. The PDFH provides monetary values of the cost that an individual attaches to travel on a crowded train. These costs vary according to the load factor, type of flow and having a seat as opposed to standing. The load factor is set at 160 per cent.

The analysis is symmetrical, that is, demand, fares, ticket shares and restrictions, generalised journey time elasticities, modal shares, and information on tastes and preferences are assumed to be the same in both directions of travel[5]. The ticket restrictions are in terms of the timing of outbound and return travel with a reduced fare ticket. The ticket restrictions are taken from the above fares manual.

The cost model is based on confidential data that has been adapted as a result of discussions with experienced current and former railway staff. There are only two elements in each cost model, a variable cost per train kilometre and a variable cost per train hour. There is no fixed element in either cost model because the empirical analysis focuses on changes in social welfare. The

[5] The analysis is symmetrical because the demand data was in this form.

variable cost per train kilometre for each operator consists of three elements, a variable track access charge, an electricity for traction charge and a maintenance cost. Each element varies according to the type of rolling stock and train length.

The electronic timetable that the SRA provided showed that Anglia Railways in the main used electric locomotive hauled coaches to meet its Wednesday 2000-2001 winter timetable obligations between Norwich and London[6]. Since very few services were run using diesel multiple units it is assumed in the empirical analysis that Anglia only uses electric locomotive hauled coaches between Norwich and London. There were only Class 86 electric locomotives and Mark II electric locomotive hauled coaches in the Anglia fleet at the time (TAS, 2002). Moreover, Anglia in October 2000 essentially operated fixed formation ten coach trains, which consisted of nine passenger coaches and a buffet car. In October 2000 the seating capacity of an Anglia Norwich-London service was 520.

In comparison the electronic timetable indicated that First Great Eastern operated Ipswich-London services with a combination of 90 mph and 100 mph electric multiple units. The First Great Eastern fleet at the time consisted of Class, 312, 315 and 321 electric multiple units (TAS, 2002). It is assumed in the practical modelling that all services are run with Class 321 stock. The First Great Eastern Ipswich-London services have always been operated with a single four car unit. In October 2000, a four car unit had a seating capacity of 231[7]. The seating capacities are also a demand model input because the capacities determine if there is any excess demand on individual services. Excess demand is recycled through the model to the next best alternative service. Demand is not reallocated to alternative modes because it is assumed that the individual has already made the choice to travel by rail in the upper nest of the hierarchical logit model. Moreover, by including Norwich on the network, spare capacity for Ipswich-London business on the Anglia services that come from Norwich can be estimated.

Finally, the variable cost per train hour captures the train crew and driver costs. The variable cost per train hour is based on a shift length of 7.5 hours. In practice, a flexible shift length system is in operation where the shift length can vary between five and eleven hours. A 7.5 hour shift length in conjunction with train hours on the above network are used to calculate the number of

[6] In actual fact, Anglia only ran two services per day from Norwich to London with diesel multiple units.

[7] Dr Tony Fowkes kindly provided the information about train lengths and seating capacities.

drivers and conductors that each operator requires to fulfil the Wednesday winter 2000 timetable. It is assumed in the empirical analysis that there is one conductor on each First Great Eastern service and two conductors on every Anglia service. The train crew on First Great Eastern and Anglia services reflects the Network South East and InterCity status of the operators.

6. EMPIRICAL METHODOLOGY

A distribution of tastes and preferences, and preferred departure times for a sample of 100[8] individuals is used to allocate passengers to individual services. The distribution is taken from MOIRA. For each individual in the sample, PRAISE estimates the generalised cost for each return service and ticket type combination. Each combination is allocated a probability in the lower nest of the hierarchical logit model. An average probability across the individuals in the sample is taken to obtain the market shares of each return service and ticket type combination. The upper nest of the hierarchical logit model is then used to determine the number of individuals that travel by rail between each origin-destination pair on the network. The markets shares are then used to assign passengers to return service and ticket type combinations.

The empirical analysis involves simulating a set of monopoly scenarios and two timetable changes in conjunction with a variety of fare combinations. The monopoly scenarios obviously involve withdrawing the operator specific tickets of the entrant. This is the only change to the base situation in the first monopoly scenario. Each of the 16 other monopoly scenarios simulates the removal of the operator specific tickets and a fare change for one of the four tickets. For each ticket a 10% and 20% fare increase and decrease are simulated. The profit maximising first class, full fare, reduced fare and seasonal fare changes are then simulated in one scenario. The two timetable changes are a higher and lower First Great Eastern off-peak service frequency. The off-peak service frequency of the entrant on the Ipswich-London flow is increased because it would have been the most likely change to the First Great Eastern service level had the SRA not chosen to create the new Greater Anglia franchise. First Great Eastern would almost certainly want to make more efficient use of rolling stock in the off-peak. As one director of First Great Eastern put it:

[8] A sample size of 100 individuals is consistent with the band of acceptable sample sizes suggested within the PRAISE software.

> 'As we all know there is no such thing as a free lunch but the off-peak gets close to it, particularly between the peaks on Monday-Friday' (Steel, 2001)[9].

From a modelling perspective, restricting the service frequency changes to the off-peak simplifies the analysis because the First Great Eastern rolling stock leasing charges do not change so they can be omitted from the First Great Eastern cost model. Even though more rolling stock is out of service in the off-peak when the lower First Great Eastern Ipswich-London service level is simulated, the rolling stock leasing charge will remain unchanged because the stock will be needed to operate services in the peak.

The higher First Great Eastern service frequency on the Ipswich-London flow is an increase in the midday off-peak service frequency in both directions from two to three services per hour. Each additional hourly service is essentially timetabled at regular intervals throughout the midday off-peak. The higher service frequency is particularly aggressive and consistent with the reputation that First Group has built up because the additional services depart shortly before an Anglia service. For example, the 0935 First Great Eastern Ipswich-London service that is introduced departs 7 minutes before an Anglia service. The lower service frequency that is simulated involves withdrawing one First Great Eastern service per hour in the midday off-peak.

The fare combinations that accompany each service frequency change are combinations of no change, and 10 and 20 per cent changes in the reduced fare market on the Ipswich-London flow. The combinations are in both directions of travel because the analysis is symmetrical. Each round is solved in the reduced fare market because there has been the most competitive activity in this market. It is still possible to assess the desirability of alternative levels of competition in the market even when the two rounds are only solved in the reduced fare market because it is assumed that everything else remains unchanged. Therefore, the effect of the service frequency changes will still be felt in the first class, full fare and seasonal fare markets. It is the effect in the above three markets of the fare changes instigated by each service frequency change that are not captured.

To find the sequential equilibrium in the reduced fare market on the Ipswich-London flow it is assumed that the operators move sequentially as they do in the theoretical model that has been

[9] Mr L. T. Steel was appointed to the First Great Eastern Board of Directors on 09/09/1998 (TAS, 2002).

outlined above. The player that adds or withdraws services from his timetable is the leader. There is only one follower in both empirical games because the games are simple two player games. The leader plays with probability δ_t the best response to fight by the follower. Fight by the follower is interpreted as a fare cut of any magnitude. It is assumed that the leader sums the pay-offs from each of his strategies across all the fighting responses by the follower. The leader plays the strategy with probability δ_t that yields the highest pay-off. Similarly, with probability (1-δ_t) the leader plays the best response to accommodate by the follower. Accommodate is taken to be a fare increase of any magnitude and the no fare change strategy. Again the leader sums the pay-offs from each strategy across all the accommodate responses by the follower. The leader with probability (1-δ_t) plays the strategy that maximises his pay-off. By the time the final stage of each round is reached, the follower Anglia will know with certainty the move of the leader. Anglia responds to each move by the leader with the strategy that maximises his long term pay-off. This may or may not involve responding with the strategy that maximises his pay-off in the round that is being played.

7. RESULTS AND ANALYSIS

For each scenario PRAISE estimates the change in social welfare. Hence, for each operator PRAISE estimates the change in profit. Direct integration is also used to estimate the change in consumer surplus for each scenario. The profit changes for each operator are commercially sensitive so the changes that are reported are in terms of the percentage change in surplus over direct cost.

The profit maximising combination of fare changes under monopoly is 20% first class fare cut, 20% full fare cut, 20% reduced fare increase and 20% seasonal fare increase. Removing the operator specific tickets in conjunction with the above fare combination results in a substantial 56.3% increase in social welfare. Producer surplus increases by 42.4% and consumer surplus increases by 11.2%.

A definitive sequential equilibrium in the lower service frequency game could not be found because the short term and long term profit objectives of the follower, Anglia, were mutually exclusive. In other words, Anglia cannot maximise profit in the present period by playing fight. A definitive sequential equilibrium cannot be found when the short and long term profit objectives of the follower are mutually exclusive because λ and ξ in equation (7) are unknown. An acceptable value for λ could be found without too much difficulty, and stated and/or revealed preference techniques could be used to estimate ξ. It is however possible to make a conjecture about the sequential equilibrium in the lower service frequency game on the basis of the investment that Anglia would commit to reputation acquisition. The conjectured sequential equilibrium in the lower service frequency game and the actual sequential equilibrium in the higher service frequency game can then be compared with the corresponding Nash-Bertrand equilibrium.

In the game when the First Great Eastern service level is cut, the Nash-Bertrand equilibrium is no fare change by First Great Eastern and a 10% fare increase by Anglia. Social welfare falls by 24.3% because there is a 6.8% fall in producer surplus and a 17.5% fall in consumer surplus. The profit of First Great Eastern falls by 4.49% and the profit of Anglia falls by 1.84%. The profit of the reference operator falls by 0.44% because the less frequent First Great Eastern London-Ipswich service reduces the number of Brighton-Ipswich journey opportunities. Consequently, some Brighton-Ipswich business switches to an alternative mode or does not travel at all.

When the First Great Eastern service level is increased, the Nash-Bertrand solution is a 20% reduced fare cut by First Great Eastern and a 20% reduced fare cut by Anglia. Social welfare increases by 32.6%. The rise in producer surplus accounts for a 5.1% increase, the remainder is due to an increase in consumer surplus. The 5.1% increase in total operator profit is the result of a 4.16% increase in First Great Eastern profit, a 0.02% increase in Anglia profit and a 0.95% increase in South Central profit.

The above two Nash-Bertrand solutions may not seem plausible because there is a marginal increase in Anglia profit in the game when a higher First Great Eastern service frequency is simulated and Anglia profit falls when the First Great Eastern service frequency is cut. The most likely explanation for what might seem at first glance counterintuitive changes in Anglia profit concerns the Anglia and First Great Eastern ticket restrictions. The Anglia and First Great Eastern tickets are not perfect substitutes because Anglia sells an inter-available set of tickets and the corresponding First Great Eastern tickets are dedicated fares. Although the new First Great Eastern services abstract some business away from Anglia this is more than offset by rail market growth. The additional services increase the number of journey opportunities, which in turn leads to an increase in the rail modal share. Some of the business that is abstracted from other modes will use an Anglia service for the outbound leg and a First Great Eastern service for the return leg, so Anglia will get some new business as a result of the higher First Great Eastern service level. For the same reason, when a lower First Great Eastern service frequency is simulated, Anglia profit falls because the number of Ipswich-London and London-Ipswich journey opportunities with an inter-available ticket falls considerably.

Attention now turns to making a conjecture about the sequential equilibrium when the lower and higher First Great Eastern service frequencies are simulated. In both rounds, First Great Eastern plays 10% reduced fare cut and no fare change with probabilities δ_t and (1-δ_t). In other words, the loss minimising strategy for First Great Eastern when Anglia plays fight and accommodate is 10% reduced fare cut and no reduced fare change, respectively.

In the game when the lower First Great Eastern service frequency is simulated, the loss minimising strategy for Anglia in the present period is 10% reduced fare increase when First Great Eastern plays 10% reduced fare cut. Social welfare falls by 24.7%. Producer surplus falls by 7.5% and consumer surplus falls by 17.2%. The fall in First Great Eastern profit accounts for 4.5% of the fall in producer surplus. Anglia profit falls by 2.58% and South Central profit falls by 0.44%. Similarly, the loss minimising strategy for Anglia if First Great Eastern plays no reduced fare change is 10% reduced fare increase. First Great Eastern, Anglia and South Central profit falls by 4.5%, 1.8% and 0.5%. Consumer surplus also falls by 17.5% and social welfare falls by 24.3%.

Moreover, although the 10% reduced fare increase is the loss minimising response for Anglia in the current period it is not the most aggressive possible response to 10% reduced fare cut by First Great Eastern. If Anglia responds with no fare change, profit falls by 0.82% from the loss minimising level in the present period. A more aggressive response would reduce profit further. Responding with 10% and 20% reduced fare cut reduces profits by 5.42% and 9.83%, respectively. Similarly, 10% reduced fare increase is not the most aggressive possible response to no reduced fare change by First Great Eastern. A response of no reduced fare change, 10% reduced fare cut and 20% reduced fare cut reduces profit by 3.9%, 13.3% and 32.3%.

The profit maximising response by Anglia in the game when the higher service frequency is simulated is 20% reduced fare cut irrespective of the fare strategy of First Great Eastern. The strategy combination 10% reduced fare cut and 20% reduced fare cut yields 32.9% increase in social welfare. The increase in producer surplus accounts for 6.4% of the rise. First Great Eastern, Anglia and South Central profit increases by 4.14%, 2.27% and 0.95%. Furthermore, social welfare rises by 33.1% when the fares combination is no fare change and 20% reduced fare cut. The increase in consumer surplus corresponds to 26.0% of the rise. The remainder of the increase is due to a 3.9%, 2.27% and 0.95% increase in First Great Eastern, Anglia and South Central profit. Responding with 20% fare cut allows Anglia to maximise profit in the present period and make the biggest possible contribution to reputation acquisition because 20% fare cut is the most aggressive strategy that is simulated. Moreover, the investment in reputation is costless.

Although Anglia is very profitable on the Norwich-London route, Anglia will almost certainly use some of this profit to cross subsidise unprofitable services elsewhere in the franchise. In addition, Shaw (2000) considers Anglia Railways to be one of the most financially challenging franchises. Moreover, the Anglia Ipswich-London inter-available fares have generally increased annually. On this basis it is argued that Anglia would only be prepared to respond with no reduced fare change when the lower First Great Eastern service frequency is simulated. When First Great Eastern plays 10% reduced fare cut and Anglia plays no reduced fare change, social welfare falls by 23.1%. There is a 7.7% fall in producer surplus because First Great Eastern, Anglia and South Central profit fall by 4.67%, 2.60% and 0.44%. On the other hand, if Anglia responds with no fare change to 10% reduced fare cut by First Great Eastern, there is 22.7% fall in social welfare. The fall in consumer surplus accounts for a 15.7% fall and the remainder of the fall is due to a 4.66%, 1.91% and 0.44% fall in First Great Eastern, Anglia and South Central profit.

8. CONCLUSIONS

The overriding feature of the various equilibrium solutions to the higher service frequency round that has been played is that more competition in the market is socially desirable when rivalry is akin to head-on competition because market growth more than offsets the abstraction of profit. Similarly, all the equilibrium solutions to the lower service frequency round suggest that market contraction more than offsets the abstraction of profit. The ORR and the SRA favour competition at the margins but the simulation results suggest that the policy makers are wrong to overlook the welfare improvements that can result from head-on competition. With competition at the margins, abstraction of profit is minimal but there is only limited prospect for market growth so the net gain will necessarily be small. The net gain from head-on competition could be significant but the bigger trade-off between abstraction and market growth will attach more risk to the profit stream of a TOC, which may in turn erode the value of a franchise. The SRA would certainly want to avoid this so a franchisee should therefore be allowed an acceptable period of time to fill gaps in the timetable before an open access entrant can enter into negotiations for the paths. Hull Trains entered because the incumbent GNER did not fill the gaps in the Hull-London timetable. At present, there are very few commercial routes on the network where there is spare capacity that an open access entrant could utilise.

There will only be spare capacity in the short to medium term if more efficient use is made of existing capacity. The SRA proposed a serious of mergers to achieve exactly this. The first merger will create the new Greater Anglia franchise. Proposals to merge other franchises are currently being considered. The results from the empirical analysis suggest that a monopoly operator will lead to a significant rise in social welfare. There will a Pareto efficiency gain because producers and consumers alike will be better off. These results however are based on no service cuts. It seems however, that the SRA will seek to preserve service levels when franchises are merged because the initial Greater Anglia Ipswich-London timetable will consist of the current Anglia and First Great Eastern services.

REFERENCES

Association of Train Operating Companies (2003): *Passenger Demand Forecasting Handbook*, London: ATOC.

Baumol, W. J. and D. F. Bradford (1970): 'Optimal Departures from Marginal Cost Pricing'. *American Economic Review*, 60, 265-283.

Baumol , W. J. (1983): 'Some Subtle Issues in Railroad Regulation'. *International Journal of Transport Economics*, 10, 342-355.

Department of the Environment, Transport and the Regions (2000): *Transport Statistics Bulletin NTS*: 1997/1999 *Update*, London: DETR.

Department of Transport (1993): *Gaining Access to the Railway Network*: *The Government's Proposals*, London: DTp.

Fowkes, T. and C. Nash (1981): *Analysing the Demand for Rail Travel*, Aldershot: Avebury.

Johnson, D. and G. Whelan (2003): 'Using PRAISE to Model Competition on Passenger Rail Networks'. First Conference on Railroad Industry Structure, Competition and Investment, Institut D'Economie Industrielle, Toulouse.

Jones, I. (2000): 'The Evolution of Policy towards On-rail Competition in Great Britain'. *Journal of Transport Economics and Policy*, 34, 371-384.

Kreps, D. M. and R. Wilson (1982a): 'Reputation and Imperfect Information'. *Journal of Economic Theory*, 27, 253-279.

Kreps, D. M. and R. Wilson (1982b): 'Sequential Equilibria'. *Econometrica,* 50, 863-894.

Nash, C. (2003): 'Rail regulation and control in Britain- where next?' 8th International Conference on Competition and Ownership in Passenger Transport, Rio de Janeiro.

Office of the Rail Regulator (1994a): *Competition for Railway Passenger Services. A Consultation Document*, London: ORR.

Office of the Rail Regulator (1994b): *Competition for Railway Passenger Services. A Policy Statement*, London: ORR.

Office of the Rail Regulator (1994c): *Railtrack's Track Access Charges for Franchised Passenger Services*: *Developing the Structure of Charges. A Policy Statement*, London: ORR.

Office of the Rail Regulator (1998): *New Service Opportunities for Passengers. A Policy Statement*, London: ORR.

Office of the Rail Regulator (2000): *Periodic Review of Railtrack's Access Charges*: *Final Conclusions*, London: ORR.

Office of the Rail Regulator (2001a): *Moderation of Competition. An Interim Statement from the Regulator*, London: ORR.

Office of the Rail Regulator (2001b): *Model Clauses for Track Access Agreements*: *Access Rights and Moderation of Competition*, London: ORR.

Office of the Rail Regulator (2003): *Changes to Access Rights and Moderation of Competition*: *Draft Conclusions*, London: ORR.

Osborne, A. (2003): 'First Group bids 500p a share for GB Railways'. *The Daily Telgraph*, 17 July 2003.

Preston, J. M., G. A. Whelan and M. R. Wardman (1999): 'An Analysis of the Potential for On-track Competition in the British Passenger Rail Industry'. *Journal of Transport Economics and Policy*, 33, 77-94.

Shaw, J (2000): *Competition, Regulation and Privatisation of British Rail*. Aldershot: Ashgate.

Steel, L. T. (2001): 'Using the Off-peak Market to Drive Profitability'. Maximising Passenger Revenue Conference, London.

Strategic Rail Authority (2002a): *The Strategic Plan*. London: SRA.

Strategic Rail Authority (2003a): *Fares Review Conclusions 2003: Britain's Railway, properly delivered.* London: SRA.

Strategic Rail Authority (2003b): *Capacity Utilisation Policy*: *Network Utilisation Strategy*. London: SRA.

Strategic Rail Authority (2003c): *Three Qualify For New Greater Anglia Franchise*, London: SRA

TAS (2002): *Rail Industry Monitor*. Skipton, Yorkshire: TAS Publications and Events.

Whelan, G. A. (2002): 'Analysing the Impact of Changes in On-rail Competition-Model Development'. Final Report to the Strategic Rail Authority, Institute for Transport Studies, University of Leeds.