

**Vertical Integration Versus Infrastructure Separation for Railroads:
Different Optimums for Different Settings?**

Alex Lu

MIT Center for Transportation Studies, 77 Massachusetts Avenue, Cambridge, MA 02139.
lexcie@zetnet.co.uk or *lexcie@mit.edu*

Word count: 7,503 words.

DRAFT 5.1.0 - 8/10/01

**SUBMITTED FOR THE 81ST ANNUAL MEETING OF THE TRANSPORTATION RESEARCH BOARD,
WASHINGTON, D.C., JANUARY 2002**

ABSTRACT

Open access is sometimes seen as promoting rail competition effectively when infrastructure geography and market demand restricts routings - e.g. in Europe. European railways are passenger-oriented, highly scheduled, poorly standardized, and lines serve specialized functions. Conversely, American railroads are freight-oriented, flexible, and highly standardized. Consequently, optimal forms of organization differ.

American railroads compete vigorously for through-traffic, and seek efficiency gains through competition and mergers. European railways are focused on local traffic, and are consolidated on a national basis. The technical, cultural, national and corporate incompatibility between European national systems precludes vertically-integrated parallel competition as a solution, requiring the operational complexity of infrastructure separation to create a pan-European network. However, American railroads, with some mandated infrastructure divestment, may compete positively, yet generate value effectively through creative inter-modal cooperation as true transportation retailers, without resorting to open access. Efficient organizational structure will ensure rail's survival through the 21st Century.

INTRODUCTION

The open access debate is taking an increasing profile in Europe and North America. In Europe, infrastructure separation is becoming all but universal. In North America, bulk and carload shippers, keen to see more competition, are calling for open or otherwise competitive access, while operators and some customers are vehemently opposed to the idea. The issue seem highly polarized. However, many factors require careful consideration by regulators and senior managers - with the ultimate goal of promoting positive competition and building a sustainable system of transportation. For any given network, perhaps there exists an optimal degree of openness, neither totally vertically-integrated nor fully infrastructure-separated.

Throughout the iron road's 175 years of history, it has shown a remarkable willingness to adapt with the times. However, railroads cannot survive if they do not continue to adapt to the transportation needs of the 21st Century. Such evolution requires an institutional structure geared to cope with the increasing pace of change. The competency of the traditional command-chain structure on a continent-wide scale is called into question, as recent large-scale service breakdowns and undervaluation of businesses demonstrate. Yet, serious errors in reorganization will result in a non-functional network, which can undermine the long term future of the railroad. The creation of a workable structure will ensure rail's survival well into the next century.

British Rail, one of the first railway systems in the world to attempt infrastructure separation, has become the subject of wide controversy as the world analyzes its effects. The geography, political need for competition, scale of the system, freight/passenger focus, historical development of the railroad, and the local institutional culture require careful focus before any restructuring. These are vital considerations in determining the optimal regulatory strategy. Comparisons between North America and Europe are particularly instructive, as they will show why the optimal forms of organization differ.

The privatization and infrastructure separation of British Rail has often been described as mistaken. The optimum degree of openness in this case actually depends on whether one takes a narrow British, passenger focused view, or a pan-European, freight dominated view. Open access may be a necessary evil to promote domestic competition - and driven by the political need to franchise operations, rather than sell them outright.

The differences in infrastructure system design, geography, and cultural environment between Europe and North America are quite startling, and are inadequately understood by some who seek to simplify the problem. Transportation is a social phenomenon, and decisions are made by customers, bureaucrats, and investors. The collective actions of the public often defy logic, and understanding this illogic is imperative to creating a functional structure that will enable sustainable transportation to develop.

DIFFERENCES IN SYSTEM DESIGN

Different railroad systems are designed to cope with different traffic patterns. This design is partly constrained by the pre-existing geographical barriers, economics, politics, and the availability of technology throughout the system's history. To determine the optimal regulatory strategy that creates a sustainable transportation system in a market economy, the design and the intended purpose of the existing infrastructure must be taken into account.

Freight versus Passenger Orientations.

Although the first railway in Britain, the Stockton & Darlington, developed to carry coal, today's rail network has a very passenger-oriented focus, typical of European railways. This is in part a political

decision by the government, since Britain's size allows reasonable journey times for express passenger trains between major population centers, while most carload freight traffic would be considered 'short-haul' - under 500km (300 miles), in a North American context. Despite the government's calls for transfer of freight from road to rail, there are few commercial incentives for shippers to choose rail. This passenger focus is partly due to system design, creating a chicken-and-egg situation. The government actively promotes rail as a real alternative to the car. High fuel duty and resulting gasoline prices at roughly \$1.32/litre (\$5.00/gallon) makes it unattractive to drive. Any freight schemes are scrutinized as to minimize their impact to timetabled passenger services, in stark contrast to North American practice.

Physical and Operational Differences.

The existing design of Britain's railways does not allow highly profitable freight services, even on important mainlines such as the West Coast Mainline (WCML), because of many operational restrictions:

- Typical sidings are approximately 642m (2,106ft) in length, making running long trains difficult. Upgrades to allow longer trains has thus far only been done on a piecemeal basis.
- Sidings are vital to freight operations in both single- and double-track territories, because bi-directional signalling has not traditionally been provided even with CTC. Sidings are required for passing trains even in the same direction. Where present, sidings and bi-directional main tracks are not used for overtaking at speed.
- Four tracks are only found in areas where mainlines are saddled with heavy suburban traffic, where they are mostly used to segregate intercity and local passenger trains.
- Many heavy-haul diesel locomotives are not designed for multiple unit operation except with the same model, and then only in emergencies.
- Freights are normally operated with just an engineer - who requires additional training to operate a train more than about 804m (2,640 ft) long.
- The draft strength of conventional European drawhook couplers limits the standard unit coal train length to 36 two-axle wagons, with a combined payload of 1,080 tonnes (1,190 tons), about 1,760 tonnes (1,940 tons) gross (1).
- Although new freight cars for heavy-haul coal and aggregates are all being built with American-style autocouplers, the older types are still in use, partly due to lack of suitable unloading/stockpiling facilities at customers' premises (2).
- New track-friendly four-axle freight cars are capable of 120km/h (75mph) rather than 96/72 km/h (60/45mph) loaded/empty for the 35+ year old two-axle types. They are connected to the locomotive by 'adapter' cars with a drophead coupler and buffers at both ends. The weak drawhook coupling between the locomotive and formation therefore limits trainloads, although some locomotives are already fitted with 'swinghead' autocouplers (3), and others can be quickly converted if facilities to run longer trains became available.
- The highest axle load generally allowed without special dispensation is 25.4 tonnes (55,800 lbs). This is due to a combination of low crosstie density, the use of 54 kg/m (113 lbs/yd) rail, and the low-tolerance alignment required for high-speed running. Only recently has 60 kg/m (125 lbs/yd) rail been adopted as renewal standard.
- High density route-signalling is installed on shared freight/suburban lines, controlled under track-circuit block regulations. The typical distance between signals is dictated by the braking capability of high-speed diesel multiple units, legally restricted to a service maximum of 0.889 ms⁻² (1.99 mph/s). At 144km/h (90mph) with four aspects, this translates to about 681~1,136m (750 to 1,250yds) between signals on level right-of-way, depending on the line capacity required. The shorter the blocks, the higher the capacity, but the shorter the maximum train length.

Under normal circumstances, the train must fit between two 'overlap' track circuits. The maximum train lengths are therefore practically limited to between 503m and 960m (1,650ft to 3,150ft) for existing installations. Even if special provisions were made in the signalling equipment, a train occupying two overlaps will still take up two train paths.

- Whilst American signalling had generally increased line capacity by increasing the number of aspects and allowing more than one train in a block, British signalling shortened blocks. Permissive working for freight following passenger had only been generally allowed since the unbraked and vacuum braked wagons were phased out in the mid-1980s.
- The W10 loading gauge, designed to carry 2,896mm x 2,438mm (9'6" x 8') containers on 945mm (3'1") high flat cars, is widely available on the WCML, but there are no definite plans to further enhance this to allow wider containers - 3,185mm x 2,600mm (10'5" x 8'6") or piggyback operation (4). Any increased clearances would incur large costs due to the need to raise the overhead wires, which have a minimum clearance of 4,165mm (13'8"). The non-electrified routes are even more restrictive due to low bridges, with an ongoing renewals programme that does not allow for double-stack operation. This compares with the standard North American freight loading gauge of 6,096mm (20') high and 3,201mm (10'6") wide outside the Amtrak Northeast Corridor (NEC).

Culture of Timetabled Operation.

In general, freight trains operate at night, similar to Amtrak's NEC practice. Freighters are permitted to run at 120km/h (75mph), unless restricted. Theoretically, each train path must be fully validated and conflicting movements checked before the train is allowed to proceed. The working timetable for freight operations (1) is more-or-less adhered to on busy mainlines. Freighters will occasionally depart early, but signalmen are instructed to prevent any early-running freighters from impinging on the timetabled paths of any other trains, freight, passenger, or deadhead moves. Restrictive trade-union conditions under British Rail giving drivers defined 'rest periods' further cemented the tradition of timetabled operation, unlike American freight operations. However, since English, Welsh & Scottish Railway (EWS) assumed control of the former BR Railfreight operations, there has been a shift towards greater flexibility.

This timetable culture maximizes the number of trains that can be operated, as conflicting moves are minimized by small adjustments to departure and point-to-point times. This results in a robust, but relatively inflexible, operating plan, which allows limited automatic dispatching. Short-term planning is done by control centers where personnel are constantly available to validate paths. Extra trains are not normally accommodated unless they do not impinge upon timetabled services. Some freight services are run as required, but their paths have been timetabled as 'reserve' paths. These may be infringed on by other extras, although timetabled extras have preference.

Geographic Differences in Network Configuration.

Because of the higher density of lines, and impact of nationalization, geographic patterns of resource utilization have developed differently from North America. For this reason, it is difficult to resurrect the vertically-integrated structure seen during the 'Grouping' era, 1923~1948.

Specialization of Mainline Designs.

Nationalization in 1948 brought parallel high-quality mainlines belonging to formerly rival companies under one roof. Where the traffic level warranted, additional resources were allocated to the line with the best business potential, regardless of heritage. The differences in characteristics

now seen between the former Glasgow & South Western (G&SW) mainline and the Caledonian Railway mainline between Glasgow and England could not be more striking. The G&SW mainline ran via Kilmarnock, Dumfries to Gretna Jct. (186km, 116 miles) where it held trackage rights to Carlisle, a strategic gateway, whilst the Caledonian mainline ran through virtually unpopulated farmland and over the challenging Summit of Beattock to Carlisle, over a shorter distance (163km, 102 miles). The G&SW sees many slow-moving through freights, particularly from the port of Hunterston, while the Caledonian handles hot intermodals and even faster passenger trains. With electric traction, the Caledonian's 1.4% grade rarely reduces train speeds to below 128km/h (80mph). Today, the Caledonian has continuously welded rail, 3 or 4-aspect CTC signalling, and a 176km/h (110mph) top speed, whilst the G&SW still has 1940's jointed rail, semaphore signals, and 88km/h (55mph) track speed, with a single line section capable of handling just three trains an hour in either direction. This traffic separation avoids the costly waste of capacity when express and heavy-haul trains are sent after one another. Although in other cases the distinction may not be so clear-cut, differential investment due to local circumstances often resulted in mainlines which are 'fit' for contrasting functions.

In terms of intermodal transloading, the key North Sea container ports of Harwich and Felixstowe are both connected to the former Great Eastern high-speed mainline to London. They are also served by a relatively slow freight line to Ely, Cambridge, and beyond, for cross-country traffic. Rail lines evolved for dedicated purposes: containers destined for the Northwest of England may enjoy shorter transit-times going the long-way-round via London, in parallel with passengers. Nonetheless, there is meaningful competition between the ports of Tilbury, Felixstowe, and Southampton. All feature mainline connexions suitable for intermodal freight, and reach the Midlands' industries independently through broadly similar infrastructure.

Parallels to this configuration manifest themselves in North America in two forms. In the Northeast where there is great speed differential between trains, 'fast' and 'slow' corridors have evolved. Between Washington, D.C. and New Jersey, where the Pennsylvania formerly competed with the Baltimore & Ohio (B&O), Reading, and Central of New Jersey (CNJ). Amtrak's NEC (ex-Pennsylvania) is a high-speed passenger route, while the B&O-Reading-CNJ route (owned by CSX and local commuter authorities) is a freight line with some commuter services. Another example is the sinuous Erie alignment between the New York area and Buffalo versus the faster New York Central (NYC) line, where the difference is largely between local and through-freight.

In low-density areas where the majority of railroads are single-track, 'direction' has become the niche exploited by formerly competing railroads. For example, between Houston, Texas, and Dexter, Missouri, Union Pacific (UP) has instituted a pair of uni-directional single lines, amalgamated from former Southern Pacific (SP) and UP trackage. Considerable productivity increases are available from such schemes, but it is important to preserve competition when the schemes are instituted.

The Rationalization of Rural Arteries, and Access to Key Freight Facilities.

The ruthless rationalization by British Rail in the 1970's and 1980's have created many "bottleneck" properties. The Caledonian Railway's Perth-Aberdeen high speed mainline competed with the slower North British route via Dundee and Arbroath prior to 1973. However, the total elimination of the Caledonian route between Stanley Jct. and Kinnaber Jct. leaves a single trunk line serving the port of Aberdeen (North East Scotland), important for its petroleum-induced passenger, timber, intermodal, and chemicals traffic. The diversity of traffic types originating from this wide catchment area made it difficult to assign the line to a single business sector. The Caledonian's reinstatement remains possible, although the traffic levels do not warrant private investment.

The Insurmountable Bottleneck Properties, and Inter-modal Interactions.

Although not exclusive to Britain, this issue is particularly acute there because of the early elimination of competition, first by grouping and then by nationalization, coupled with the demand for shortest technically-feasible journey time from the time-sensitive logistics and passenger businesses. For these sectors, the only realistic competition with the Severn Tunnel in the West of England is an expressway bridge; the only competition with the Forth Bridge is a highway bridge; and the only effective competition with the Channel Tunnel are the high speed ferries and air.

These properties are highly strategic, and represent the only reasonable direct rail routes across major geographical barriers. Due to the relative proximity of important city-pairs, the small time differential becomes very important, as illustrated by rail's clear majority market-share for passenger travel between Edinburgh and London, despite much cheaper bus alternatives. The market is so sensitive to journey times that if 80km/h (50mph) speed governors were fitted to high-gross trucks, the whole economics of logistics would be severely disrupted.

Interestingly, the strategic importance of these properties decreases in a broader European context, because the significance of the two to four hour differential falls as overall shipping time increases, especially for freight. Wider varieties of routes that deliver to broadly similar areas become competitive, yielding a picture similar to that in North America. For intermodal freight going from Berlin to Glasgow, the Channel Tunnel is an all-rail solution with a long detour, while a ferry crossing from Hoek van Holland to Hull is probably more cost-effective. Likewise, times would be comparable for docking a transatlantic container ship at Newport, South Wales, and routeing intermodally using the Severn Tunnel, and for docking at Southampton for London. Frequently, the whereabouts of docking is dictated by port facilities and levies as opposed to rail speed or cost. It is usually cheaper to dock on mainland Europe than to use the Tunnel.

For bulk commodities, the preferred method of cross-channel shipping usually involves transloading. Relatively long journey times by train-ferry and the resulting poor freight car utilization mean that it is no longer a realistic proposition for cross-channel bulk freight.

In North America, distances are broadly similar to a pan-European perspective, although there is less demand for short absolute journey time between urban centers, due to the paucity of intercity passenger service. In the intermodal sector, reliability and predictability, rather than the journey time, are more important. Alternative intermodal routeings in North America commonly increase the journey time but only by some 10%, compared to at least 30% in the domestic British cases cited. Nonetheless, some partial bottlenecks exist, such as the Cajon and Tehachapi passes in Southern California, or the Hudson River at Albany.

For heavy hauls, it is almost unknown for any important city-pairs on either side of the Atlantic not to be linked by alternative routeings. Where the journey time is of little consequence, paths on Britain's rail network, particularly the Channel Tunnel, are too precious to be wasted for such applications. Due to the proximity of major European cities to the coast, much more freight moves by ship than in North America.

In Japan, where Japan National Railway (JNR) was privatized in 1987, the geography and scale are both somewhat similar to Britain. The passenger focus is reflected in the ownership, as freight operators are tenants to vertically integrated passenger railroads and account for only 5% of total ton-miles of freight moved (5). Most freight moves by ship (45%) and road (50%). Railfreight has only been able to compete for long-distance container flows. Specialization of infrastructure designs is even more polarized as Shinkansen high-speed lines carry only specialized rolling stock. Bottleneck properties are common in the form of strategic bridges, urban trackage, and city center stations. Partial nationalization since 1909 also meant that there are virtually no duplicated conventional interurban trackage, although Shinkansen lines often run parallel to conventional

lines. Japan is unique as transcontinental freight can only move by ship and air, therefore there is no larger context.

INSTITUTIONAL ORGANIZATION

All railroads were vertically-integrated at first, because it seemed logical when the lines were sparse and poorly connected. It simply made sense for the same company to own track and rolling stock, especially when there were few interchanges and sometimes even breaks of track gauge (Britain's Great Western, or the Erie in America). However, as railroads grew, institutional changes occurred as railroads reorganized themselves under stress, or were reorganized by the government due to bankruptcy. The cumulative effects, influenced largely by politics of the period, gave rise to two very different systems based on two different philosophies. North American railroads tend to focus on through traffic while European national systems are more locally or at best regionally-oriented. The North American rail policy has long been to seek efficiency through competition among more or less equal systems, while the more recent European approach has been to seek efficiency through concentrating certain types of traffic on certain lines.

Historical Development of North American Railroads.

In North America, particularly in the West, the railroads grew up to push back the frontier of settlement. Arduous and dangerous journeys which formerly took months were reduced to a week's travel in relative comfort. Generous land grants were given to the Union Pacific (UP) to generate resources for the first transcontinental mainline, which formula was successfully followed for many other railroads. In most cases, alignments were chosen on engineering rather than political grounds, creating towns en-route with little local interference. The purpose of these lines was to bond a large and diverse nation together; to enable remote areas to be settled and to encourage commerce. Indeed, the promise of the Canadian Pacific was instrumental in convincing British Columbia to be part of Canada.

Of course, a single railroad was an absolute monopoly at a time when rail travel was the only realistic way to get across the continent. One mainline was a rather large "bottleneck" property and clearly insufficient. To promote competition, and to enable wider settlement, North American federal governments encouraged several other land-grant railroads to be built over broadly parallel alignments. The culture of competition was inbuilt during the construction race, and has remained with railroaders ever since. The railroads competed head-to-head for lucrative long-distance traffic, while retaining sufficient 'captive' traffic to remain profitable. In an era when the core revenue was generated from local carload traffic, these captive customers enjoyed good service.

Farsighted promoters looked across the whole continent, as evidenced by the 'Pacific' name in many of the western railroad systems. In the buoyant 1920's, no single system was able to cope with the booming travel demand, and the combined capacities of a number of rival vertically-integrated concerns went some ways towards alleviating the problem. This was the American Railroad at its best.

In the Northeastern states, however, railroads were originally built more on a British-European model - to connect pre-existing townships. This emphasis on local rather than through traffic is evident in that a number of NYC subsidiaries remained independent, and only linked up as a 'system' on a piecemeal basis. Initially, state canal and turnpike interests interfered with both the routing and the tariffs, although neither remained influential for long. Later, state legislatures (especially in Pennsylvania and Maryland) sought to promote "their own" railroads (i.e. PRR, B&O respectively) for local business interests, although the unofficial protectionism never reached European levels. Furthermore, the eastern states were unable to protect their railroads from

competition west of the Alleghenies, as the eastern roads reached the Chicago and St. Louis gateways.

With no quasi-protectionist pasts to look back on, the western railroads focused on their mission to build transcontinental systems, and reached agreements to share resources, such as running rights over Raton Pass between Santa Fe and D&RGW, in order to reach different territories, as opposed to preventing rivals from building a system. The sheer volume of captive and induced local traffic available in the west rendered legislated protection unnecessary.

Trackage rights, leases, and union stations developed as promoters realized that cooperative competition could safeguard the future for both concerns while they accomplished the final goal of developing a profitable system of transportation. The joint development of Washington Union Station by B&O and PRR shows that, given the right legislative environment, even old foes could work together for the best interest of the transportation industry, without having to come under joint ownership or undermine competitive interests of either concern elsewhere. Rivals similarly came together to build union stations in Chicago, Cleveland, and several other cities. The Northern Pacific and the Great Northern even shared control of friendly connexions at either end - Chicago, Burlington & Quincy, and the Spokane, Portland & Seattle.

It was significant that the short period of nationalization of American railroads during World War I ended in failure. All the perils of nationalization were exposed at once then; operation was messy, labour relations became very difficult, and the system was unmanageably large. Before any significant rationalization or investment could take place, the system was returned to its private owners. Although the lack of progress during nationalization was due more to the wartime demands placed on the railway than the inherent weakness of such a set-up, nationalization and service breakdown were forever associated in the minds of leading railroaders and shippers. This experience largely silenced calls for nationalization, although recently certain mergers in the private sector have produced neo-national systems and re-introduced the service failure issue. Subsequent government ownership was limited to Canadian National and Conrail, both created to rescue bankrupt railroads and eventually returned to the private sector as prosperous systems.

Since the Depression, the dynamics of transportation had been turning slowly against the railroads. The railroads were no longer able to rely on the captive carload freight customers for their core revenue, or to generate the much-needed cash for modernization work after World War II. The only real competitive advantage remaining was rail's economies of scale over huge distances and for bulk loads. Nonetheless, railroads were experiencing surplus capacity everywhere, and signs of over-competition was apparent, especially in the high-density Northeast. This led to a rush to merge and rationalize surplus capacity (6). While Europe chose nationalization on a grand scale (sometimes with deficits to match), the North American market-economy produced mergers and bankruptcy.

Merging Systems in North America.

Two types of mergers may be distinguished: expansion-driven, and consolidation-driven. Geography is of the utmost importance when analyzing any merger. Both types of mergers can be beneficial to the transportation industry, provided that the system remains manageable and that the merger does not create an anti-competitive situation.

Expansion Driven Merger.

These mergers are peppered throughout railroad history on both sides of the Atlantic. By the beginning of the 1900s, the American railroad system was largely complete. However, it was still possible to expand by merging with neighbouring railroads making an end-to-end connexion, a trend

that continued right up to the mid-1990s. Examples include the purchases of Western Pacific (WP) and Chicago & North Western by UP, which created a transcontinental mega-system stretching from Chicago to San Francisco. Mergers of this type, like arrangements for through-running power, and trackage-right exchange schemes, reduce the number of interchanges required to ship from one city to another. Not only do they lead to better utilization of resources (power, cars, maintenance facilities, personnel), they also lead to expanded marketing and other corporate activities. Again, to the extent that the merged railroads were already acting as friendly connexions, these mergers do not pose anti-competitive threats. On the contrary, they encourage other parallel railroads to do the same and therefore boost their operating efficiency. Mergers of this nature are, in general, good for the companies, the customers, and transportation network as a whole, provided that the system does not become unmanageably huge. Indeed, they may enhance the competing power of the merged company by offering additional through-traffic journey opportunities, so that the whole is greater than the sum of its parts. In Britain, this was realized in the late 1980s as passenger locals were combined to form through cross-country "Regional Railways" services, increasing assets utilization and reducing transfers.

Consolidation Driven Merger.

Mergers of this nature cause more concern, and are more likely to occur in a declining industry seeking to cut operating costs than a buoyant one seeking to expand. These mergers are beneficial in an area where there is over competition, such as during the late 1960's in the Northeast, before the formation of Conrail. Traditionally these have been disallowed by the Interstate Commerce Commission, predecessor of today's Surface Transportation Board, such as the proposed Santa Fe/SP merger in the mid-1980s. However, where a merger had been allowed without the lines having 100% end-to-end geographies, an element of consolidation would have occurred. The regulatory question when examining such mergers is, therefore, whether the regional market showed excess capacity and clear efficiency benefits, or if the consolidation was simply an attempt to create a localized monopoly.

A few examples will illustrate. Following the failure of Penn Central, the Northeast's railroads clearly needed restructuring. Federal intervention created Conrail in 1976, and the most efficient through-routeings were chosen to create a megasystem. This is evidenced by the shift from the PRR Philadelphia-Harrisburg mainline (now Amtrak-owned and passenger oriented) in favour of the more direct ex-Reading/Lehigh Valley/CNJ line via Allentown. Although it created a near-monopoly in much of the Northeast, this was necessary to support the declining amount of originating local traffic. However, efficiency can also be lost in this sort of reorganization, as Conrail's takeover of the Reading terminated a friendly interchange with the B&O (now CSX) at Philadelphia, to divert traffic onto its own lines. Meanwhile, Conrail's focus shifted towards competition for through-traffic, where its success was clearly demonstrated in its 24 years' of profitable operation. Some competition was preserved since CSX and Norfolk Southern were able to route independently to Chicago. Fortunately, Conrail had a strong management team, which managed to carve out a niche for itself to ensure survival with expedited intermodal trains. Coupled with the Santa Fe premium service to the west coast, the 'land-bridge' concept first developed under NYC was re-born, offering a faster route than the Panama Canal, and filled surplus capacity for both carriers.

In the West, the merger between UP and SP in 1996 eliminated competition between Salt Lake City and Oakland. Since UP already held Rio Grande (D&RGW), it now held a near-monopoly on traffic between the Missouri River and Northern California. This is a dangerous situation on two grounds: the shippers now have less choice, and UP has less incentive to streamline its operations in that area. Even with trackage rights on UP between Denver and San Francisco, BNSF is still

unable to compete on a level playing field, because all lines are dispatched by UP. Had UP been ordered to sell parallel WP or SP, and D&RGW or Overland Route trackage in the area to BNSF or a third party, along with a reasonable feeder network to generate originating traffic, there might have been greater efficiency gains. Parallel one-way operations such as that between Wells and Winnemucca, Nevada, could have remained with trackage-rights and crew-sharing agreements, as it was when WP and SP were separate railroads. The UP and Missouri Pacific merger also removed much competition in the West.

Although the Burlington Northern/Santa Fe merger was more of an expansion merger than a consolidation one, there is still a small element of competition reduction. For example, BN and Santa Fe held competing alignments between Chicago and Kansas City. However, BN's core business was in the Northwest, hauling coal and grain, while Santa Fe's was in the Southwest, hauling intermodals and other port traffic. This makes it for the most part, an end-to-end merger. Crucially, before they merged, BN & Santa Fe were already jointly marketing intermodal services, based on the extended geographical reach of their networks. The merger then combined their respective expertise in heavy haul and intermodal traffic, expanding the geographical reach of two types of drastically different businesses (7). To some extent, BNSF uses the BN and Santa Fe alignments to separate traffic between Chicago and Kansas City into fast and slow lanes, and has not abandoned either line.

Historical Development of European Railways.

The railways grew up the same way they did in the American Northeast, with the routeing basically dictated by existing urban centers, although further restricted by national boundaries. Unlike North America where people and goods could move freely across state and provincial boundaries without hindrance, it was not until the Single European Act (1987) when duties were abolished in practice (8) within the European Community (EC).

Beginning with the railway mania of the 1850's until the end of Britain's Grouping era, there was widespread competition by companies with parallel routes, but this did not occur on a pan-European scale. Competition was rife for such lucrative flows as London to Glasgow and Edinburgh; also for 'boat train' flows from Exeter to London. However, these railways were cut off from the European mainland, and only competed on short-distance flows. Even before nationalization, companies developed their own specialized markets, for example, LNER had always been more focused on passenger flows than the more freight-oriented LMS. The nationalization of British railways was partly to redress such imbalances.

The focus on local traffic, and the lack of foresight, was surprising even when compared with early American railroads, mainly due to the European tradition of organic civilization growth. Part of the future West Coast Mainline was constructed as the Lancaster & Preston Junction Railway (9), which terminated in Lancaster with a stub-end, with no convenient way north! Although there is now a through connection, a 70mph speed restriction remains in place as a permanent reminder of that strategic oversight. This was typical throughout Western Europe's industrial heartlands - unlike the Santa Fe's historic efforts to secure a good alignment to Los Angeles.

On mainland Europe, national railways developed different and incompatible systems of handling traffic. Run-through services were generally disallowed on political grounds - the Rhine was as effective a barrier as the English Channel! Because of the higher density of population, industry and the national self-sufficiency culture in Europe, little thought was given to pan-European transportation. This philosophy is reflected in the widespread existence of government-backed 'flag-carrier' airlines in Europe until recently. While interstate commerce has always been encouraged in the United States, slogans such as 'Buy British - support our economy' were common as late as the 1970s. Although France, Germany and Italy had showed an interest in fostering

commerce jointly since the late 1950s, this was mainly limited to sharing natural resources. For manufacturing, each nation-state was self-interested, thus legislation was diplomatic, and implementation superficial. Companies with facilities in more than one country are relatively recent, and began only as the Tiger Economies of the Far East invested in Europe. Therefore, until recently, there was little point in having compatible railroad systems which spanned more than one country, as coal and steel travelled in barges. Although certain regions, such as Holland and the Alsace region of France, were more open to the idea of international trade, this was a very localized phenomenon.

Within Britain, the former 'Big Four' were very proud of their individual heritages and refused to accept innovations from other railways. The cultures of national railways in mainland Europe had been similarly inwards-looking. The nationalization movement was partly an attempt to harmonize the different equipment within countries - something which didn't always happen, as the BR's Western Region chose diesel-hydraulic traction to replace steam instead of the popular diesel-electric.

Although by the 1980s, British Rail was on the whole one standardized system, nationalization made the differences in railroad technology between countries far more acute. BR's Southern Region, facing the English Channel opposite France, was electrified at 750V DC third-rail, whilst BR's London Midland Region and the French chose overhead 25kV 50Hz; the Dutch went for overhead 1.5kV DC, whilst the German standard 15kV 16²/₃ Hz system was used in Switzerland, Sweden, Austria; but Italy and Belgium had a 3kV DC system. This is not to mention the differences in loading gauge, signalling, brakes, couplings, and even track gauge (for Ireland, Spain, and Finland).

The Treaty of Rome established the European Economic Community (EEC), but in retrospect it was an economic collaboration more than a genuine attempt at European integration. Meanwhile, suffering from increasing losses, the national railways looked inwards for efficiency gains by segregating traffic, rationalization, and other consolidating activities, much as Conrail did. Significantly, the barriers arising from national boundaries tilted the economics of freight against rail, and favoured development of lucrative short-haul passenger traffic under about 500km (300 miles).

Development of a pan-European Transportation System.

Despite the provision for interchange of road vehicles early on, not until the Single European Act was serious thought given to a pan-European multi-modal transportation system. The EEC abolished trade barriers amongst European nations, but did not create a cultural basis to encourage industrial consolidations across former national boundaries. More recently under the European Union (EU), attempts were made to harmonize economic development and create an European identity, but it was infested with problems due to short-term protectionism. For example, the Channel Tunnel suffered some 70 years' stop-go, despite continued support from the French, before it was finally built. High-speed international passenger services did not become commonplace until the 1990s, even on the European mainland.

Within the EU, national governments retain substantial financial independence, and each seeks to protect the interests of its own people against the rest of Europe. This protectionism is further exacerbated by the strong national identities resulting from a diversity of history, language, and differences in culture amongst Europeans. When commercial vehicles were permitted to make deliveries across Europe, there was considerable outcry in Southern England due to differences in motor vehicle legislation and fuel duty in France, whereas the difference in taxation between American states is accepted as a matter-of-fact. In North America, states compete for residents, whilst in Europe, nation-states compete for privileges. The recent clashes over mad cow disease,

North Sea fishing quotas, etc. further highlight the failure of EU citizens to identify with each other as 'Europeans' like Americans do. These underlying nationalisms within the EU, combined with the existing monopolizing national institutional structure typical of European railways, makes it extremely difficult to achieve the 'level playing field' for transportation that politicians ostensibly aspire to. Although theoretically many Continental Europeans are open to the idea of a combined nation, in practice they are not ready to give up the political protection their home states provide.

If the EU truly aspires to unity, the reorganization of its railways on a wholesale scale is necessary. Instead of small regional systems focused on local traffic and terminating at or near former national boundaries, the railroads need to form alliances like American end-to-end mergers. However, since the existing institutions are nationalized railways, which hold every important mainline, some rationalization and reallocation of geographies is required. In the American Northeast, this reorganization occurred under Conrail after the Penn Central bankruptcy. But since the current European railways are government-backed, and heavily subsidized, they are unlikely to go bankrupt. National governments, moreover, have strong attachments to their railway networks, and are unwilling to relinquish control even to achieve a more efficient and innovative pan-European system. To comply with the EU philosophy of establishing a competitive level playing field, open access was instituted to provide some competition, and to pave the way for possible future geographical reorganizations.

Harmonization of operating practices, safety systems, and the infrastructure cannot all be done at once, and, as was shown on BR, open-access is a way of encouraging this much-needed standardization work. Taking a long term view, the European Council considered mandated open access throughout Europe as the first step towards a pan-European network of transportation. The subsidiarity principle of the EU states that decisions should be taken at the lowest level possible. This is an awkward compromise reached as each nation attempts to guard its own interests while still taking strategic direction from the EU in principle. Since 1987, the EU has theoretically had the power to break up the national monopolies, but in practice there is insufficient support even amongst forward-looking politicians to implement a radical reorganization of railways. National pride is one reason, but member governments are also keen to safeguard the interests of domestic intercity and commuter passenger traffic against trans-European freight flows. Furthermore, merging two roughly equal but different systems (such as SNCF and DB) might lead to irreconcilable human conflicts as demonstrated under Penn Central (10). A more subtle approach is needed, with nationalized monopolies retaining dispatching control, while allowing competition by foreign carriers. This allows the decisions to privatize railways to be made locally, while ensuring the resulting structure meets pan-European transportation needs. The motive to institute open access really is more of a political one than a commercial one (11).

CONCLUSIONS - DIRECTIONS FOR THE FUTURE

As the highways become ever more congested, the economics of carload freight is tipping once again in rail's favor. Provided the economy remains buoyant, with increased logistical demands from e-commerce, the railroad's ability to provide bulk transportation will become more important. However, the current geographical configuration of the North American mega-systems means that the class I's are, on the whole, unwilling (or unable) to compete with deregulated truckers. In other words, the system is not optimized for carload freight. Decades of rationalization and consolidation have bred managers who are firmly focused on bulk and capacity, who like to combine or discontinue shorter trains to cut costs, while eliminating the "competition" of other railroads. While attempting to amass even greater bulk of the rail market share, the net reduction in rail's share of the total market is rarely noticed as services to small customers are sacrificed for operating efficiency (12).

An increasing number of shippers are calling for 'open-access', hoping that better service would result.

This drive to eliminate competition from other railroads is deep rooted in its cause, and not entirely the fault of railroad managers. Past regulatory policy had been to promote intermodal competition - head-to-head competition between canals, railroads, and truckers. The Pennsylvania's attempt to stave off bankruptcy with its own trucking arm was deemed anti-competitive, whilst its merger with the largely parallel NYC was not. The intrinsic difference between vertically-integrated railroads and infrastructure-separated highways, plus the importance of geography, was not appreciated. The same mistake is still being repeated in Britain, with the mandated sale of bus operator Scottish Citylink when its parent company was awarded the ScotRail passenger franchise.

In Europe, infrastructure separation appears to have been implemented as a last resort to generate some competition in freight transportation in the face of stiff opposition by EU-member governments. In Britain, where infrastructure separation was embraced early on, it was the result of domestic politics. Even without such political barriers, true vertical integration would still be difficult to implement as Europe's passenger-focused railways do not lend themselves to the creation of roughly equal competing freight operations. The demand for ever-shorter journey times between very specific urban centers on short-haul intercity service and narrow commuter corridors mean that a de-facto monopoly would exist under total vertical integration.

In North America, no such political resistance exists, and the railroad geography is such that extensive competing parallel systems existed at one time. With creative intermodal integration, any railroad should be able to reach any customer through a short-haul trucking network. The relatively greater separation between urban centers mean the delay in transloading is less significant. The service failure issues may be resolved without resorting to open access, with all the associated operational difficulties as clearly demonstrated on the former British Rail network (13).

For bulk and chemical shipments, which are difficult to transload, a nationwide realignment of institutional boundaries to create competing parallel mainlines, with federally mandated terminal trackage rights for bulk commodities only, may be just the wake-up call the industry needs to address the service failure issue. Canada's government already requires terminal trackage rights for all freight.

As long as North America's railroads remain freight-oriented and focused on through traffic, effective regional rail monopolies will continue to exist, much to the disappointment of some shippers. In the freight sector at least, the effect can be mitigated, but in the passenger sector where transferring is undesirable and delays have a larger impact, perhaps legislated protection is necessary if passenger rail is to remain viable. Future research should be focused not on 'how to make open access work' but instead 'how to promote competition' without introducing an additional level of operational complexity - perhaps by reassignment of geographies and promoting intermodal cooperation. Railroads could become multi-modal transportation companies (14). Transportation retailing may be a tough market to be in, but railroads need to learn to do what truckers have been doing since the 1930's - to add value to their business by providing reliable point to point transportation.

ACKNOWLEDGEMENTS

The author is grateful to John G. Allen, Senior Transit Analyst, Regional Transportation Authority, Chicago; Thomas Hoerber, Centre for European Studies, University of Cambridge, England; and others for their useful comments. All errors and opinions are, of course, those of the author.

REFERENCES

- (1) Railtrack Scotland. *Working Timetable, Section GF*, Glasgow, 1999.
- (2) Shannon, P. Power point! Fifoots fuels EWS traffic, *RAIL* No.400, Jan. 2001, p.28.
- (3) Dunn, P. Class 66 - a new level of reliability, *RAIL* No.401, Jan. 2001, p.40.
- (4) Railtrack HQ. *Network Management Statement for Great Britain*, London, 2001.
- (5) Japan Ministry of Land, Infrastructure and Transportation, *Annual Report of Transport Economy*, 1996.
- (6) Wilner, F. *Railroad Mergers: History, Analysis, Insight*, Simmons-Boardman, Omaha, 1997.
- (7) Frailey, F.W. Watch Rob Run, *TRAINS* Vol.61 No.1, Jan. 2001, p.60-67.
- (8) Weidenfeld W., Wessels W. (eds.) *Europa von A bis Z*, German Political Institute, Köln, 2000.
- (9) Nock, O.S. *Electric Euston to Glasgow*, Ian Allan, London, 1975.
- (10) Daughen J., Binzen P. *The Wreck of the Penn Central*, Little & Brown, Boston, 1971.
- (11) Posner, H. Global strategies for railroading. *TRAINS* Vol.61 No.4, Apr. 2001.
- (12) Hemphill, M.W. Architects of the new age, *TRAINS* Vol.61 No.1, Jan. 2001, p.42-45.
- (13) Green C. Phoenix from the ashes - rediscovering the business-led railway, *Sir Robert Reid Lecture to the Institute of Transport and Logistics*, Feb. 2001.
- (14) Smathers G. A. *The ASTRO Report*, The Association of American Railroads, Washington D.C., 1970.