THE CONTESTABILITY OF URBAN PASSENGER TRANSPORT MARKETS: MARKET SIZE AND MARKET DENSITY

J Brander and B Cook
Department of Economics
University of New Brunswick
Fredericton, NB, CANADA, E3C 5A3

1.0 INTRODUCTION

This paper examines the contestability of urban passenger markets. We examine evidence from the airline industry, drawing on that experience to help to define the attributes of industry contestability in the context of the service sector. The lessons learned are then applied to the urban bus market.

Some conventional wisdoms concerning the contestability of the US deregulated airline industry are reviewed in Section 2. The extent to which these are applicable to surface transport assessed. We next consider the issue of market congestion and other problems created by the removal of property rights (and establishing ultra free entry). This includes a discussion of the applicability of the model to urban passenger markets. Section 4 presents a formal model describing the outcome of a deregulated urban passenger industry, an outcome involving excess capacity. The outcome is inevitable, and will result regardless of market size or density. A final section of the paper draws conclusions from the analysis.

2.0 IMPERFECT CONTESTABILITY : CONVENTIONAL WISDOMS

Baumol [2] originally suggested that deregulated airline markets were fully contestable. By 1986, Baumol and Willig [3] observed that the early deregulation experience revealed supply side constraints which precluded full contestability. They enumerate problems of airport access, air traffic control, equipment availability, and wage differentials between new and old carriers. They conclude that competitive forces have been largely successful in protecting the interests of the passenger. In their view, all the impediments to contestability are on the supply side.

In discussing contestability, one must distinguish between two situations. The first is the scenario in which an existing firm, having all of the requisite facilities at both ends of a city-pair link, wishes to contest that particular market. Aside from an advertising campaign designed to attract passengers to the new service, the sunk costs of entry are zero if one abstracts from the airline using its airport capacity in one fashion rather than another. This case can be thought of as market contestability. The second case is that of a new entrant.

Here, there are a variety of sunk costs such as those of obtaining internal airport space and in certain instances, landing rights as well. This case may be thought of as industry contestability. It is important to distinguish between the two.

Shepherd [15] notes confusion over the definition of the market, and concludes that the established concepts of market structure and entry' adequately explain airline competition. He holds that the changes in market shares which have occurred come mainly from interactions among competing firms. He also appears confused about the concept of a "market". We are left uncertain as to where the market stops and the industry begins. It is perhaps this confusion which leads him to suggest that it is difficult to tell which firms are in the market and which are potential competitors.

Closely related are those actions taken by incumbent carriers to product differentiation. They can be expected to act defensively, taking actions to protect their market shares against the pressures of the new competition. A variety of actions are possible. In the airline case,
these would include frequent flyer programs and the establishment of fortress hubs. Both of these actions are readily observable, and both represent normal competitive activity. Such activities are designed to protect market share. These actions relate to issues of the share of the total industry market held by an individual carrier and not to the contestability of a particular market. Issues of the efficacy of contestability as an instrument for removing or reducing monopoly rents relate to the specific city-pair market. At this level, facility accessibility dominates product differentiation as a barrier to contestability.

Bailey and Williams [1] argue that "local monopoly rents reflect the benefits of sunk costs at a strategically located facility". They argue that the rents arise because of the ability to develop a hub and spoke network. The rents are more general. What is crucial is control over the critical groundside and airside facilities at hub airports. The rents are not of the intramarginal variety arising because of the greater efficiency of particular carriers at particular locations. Where serious airport congestion exists, the possession of these facilities generates economic rent regardless of the network configuration involved. Finally, they suggest that the whole process of deregulation was based on the expansion of airports as required by traffic growth and the active pursuit of antitrust policies. Apparently, neither occurred. In this regard, Eads [11] argues that airport access is the most important factor in determining the effectiveness of potential entry in modifying incumbent behaviour.

With reference to the airlines, a new entrant must incur costs which are not borne by the incumbents. Such costs relate largely to the acquisition of airport facilities. Ceteris paribus, the costs of a new entrant will be higher than those of an incumbent. In turn, this will have two effects. Firstly, it will reduce contestability by tilting the playing field in favour of existing carriers. Secondly, since these costs are now a normal business expense, in the long-run they will become embedded costs of carrying on business, introducing new rigidities into the industry. In the extreme, they can lead to new industry demands for entry restrictions. Both create those new deadweight losses which Tullock [17] hoped this generation would not bestow on the next.

2.1 The Question of Sunk Costs:

One of the requirements of airline contestability is the possession of the appropriate airport facilities at BOTH of the airports on the link which the entrant wants to contest. Otherwise, the entrant will have to incur sunk costs of entry. This negatively impacts on the contestability of the market, resulting in the ability of the incumbents to earn some monopoly rent. While most attention in the literature has focused upon the sunk costs which must be incurred by new entrants, the sunk costs of the incumbent carriers must also be considered. The impact of sunk costs upon contestability is the net result when these are offset against one another. Therefore, the balance could (theoretically) tilt either way, favouring either the incumbent or the potential entrant. At least this would be the case in those situations where public policy was directed at ensuring a "level playing field". It follows that in discussing the contestability of a given market, some attention must be devoted to the openness of the institutional framework in which the industry functions.

Coursey, Isaac, Luke and Smith [9] have considered contestability in the presence of sunk costs, defining them as costs which would be escaped by not entering a particular market. In their analysis, entry permits (valid for five periods) were required, their cost representing the sunk entry costs. They concluded that while sunk costs weaken contestability, they do not destroy its market disciplining power, suggesting that a weaker version of the hypothesis remains valid. Artificial entry barriers thus permit incumbents to earn monopoly rents even in deregulated industries.

Incumbents may encounter sunk costs which new entrants into the industry (not merely into a particular market in the industry) do not face during the transition to deregulation. Meyer and Tye [13] note several potential hangovers from a regulatory period. They specify choice of aircraft, labour contracts and excess capacity among other residual effects of the regulatory period. All impose sunk costs on incumbents, at least in the short-run. As a result, on balance, they conclude that "individual prices seemed to have little to do with the costs of individual services".

2.2 Implications for Surface Transport :

What are the lessons from the above for the contestability of urban passenger transport? There are three impediments to contestability in airline markets:

- infrastructure access
- product differentiation and
- sunk costs.

Each must be discussed in turn.

With regard to the first, there is a significant difference between the mechanics of operation of the two modes. In the airline case, the operations are like a pulse, a situation in which a relatively short piece of infrastructure (the runway) is utilised perhaps as many as 40 times per hour. With surface transport, the analogy would be to a continuous flow along a loop. In the worst case scenario, vehicles would queue, and while one might impede another briefly, there are no property rights issues involved. (Here we abstract from firm-owned terminals.) We conclude, therefore, that infrastructure access problems do not constitute an impediment to contestability in the case of surface urban passenger transport.

Closely related is the issue of sunk costs. Since infrastructure problems do not exist, this dimension of sunk costs, so important in the airline case, is not relevant. The Meyer and Tye proposition noted above is important here, and might well be critical for the incumbent if the issue were the privatisation of the service. For the potential entrant, sunk costs would take the form of vehicle acquisition and a variety of costs associated with attempting market entry and increasing market share. Such sunk costs are of unknown amount, though they place the entrant at a competitive disadvantage. What is interesting here is that the two aspects would partially off-set one another. The extent to which these sunk costs impede contestability in the case of surface urban passenger transport is a matter for specific case studies. Generalisation does not appear to be possible. Nor does it appear to us that there is much potential to utilise product differentiation as a means of reducing the contestability of these markets. We say this because the kinds of bonding techniques employed in the airline industry seem to have little applicability.

None of the issues discussed above constitute serious barriers for the contestability of urban surface passenger transport. While the case of ultra free entry may not apply, the threat of potential entry would be sufficiently strong as to modify incumbent behaviour.

3.0 IMPERFECT CONTESTABILITY: MARKET CONGESTION

In an economic context the major problems of deregulation stem from the external costs which the policy generates because property rights are unspecified. According to Pigou [14], the externalities created by the mutual exploitation of such a market are usually characterised by "direct interaction" of economic units within a legal framework which does not compel compensation

for disservices. This approach assumes a technological interdependence among users which persists because of a lack of specification of property rights.

Consider the case of a deregulated airline industry as an example. Where the industry is unconstrained by regulation or airport capacity, there will be increases in the number of flights until all market rents are dissipated. This rent dissipation results from the excess capacity generated by new entry and/or the expansion by existing firms. The excess capacity creates a form of technological externality resulting from the direct interaction of the airlines competing for the same passengers. Since passengers captured by one airline are unavailable to others in that time period, airlines exploit the available passenger stock in a suboptimal manner in that marginal opportunity costs are equated with the value of the industry's average product rather than the value of its marginal product. Such competitive strategies as head-to-head scheduling develop, leading to greater capacity levels and lower associated productivity than when property rights are specified. This adverse interaction among competitors is extremely important in generating market congestion. In turn, market congestion has implications for contestability because of its impact on cost structures. Before this issue is discussed, it is necessary to focus on the economic analysis of the service sector.

3.1 Some Thoughts on the Service Sector:

In neoclassical economics, the firm combines two inputs (labor and capital) to produce a good. Both inputs are required. Output would be zero if either were absent. In the long run both are variable and under the control of the firm which can therefore determine ex ante its input mix and hence its output. Markets are assumed to exist and to function in an orderly manner so that there are no transactions costs. From this emerges the traditional neoclassical outcome : constant costs, zero economic profit, and maximum consumer surplus.

The service sector is different, and so the strict neoclassical model does not apply fully. As we have argued previously [5], the firm must combine three inputs to generate its output. The service sector firm requires a third input -- customers. Sales cannot occur without customers and output and sales are not only synonymous, but also simultaneous. Inventories are not possible and the customer is an integral part of the productive process. A major implication is that the firm (since it controls only capital and labour) can only determine capacity ex ante. Output is determined ex post after the third factor of production has appeared to play its role. The firm attempts to attract customers by manipulating certain customer-specific selling costs, by altering the quality and/or frequency of service delivery, or by developing strategies designed to strengthen the bond between it and its customers.

When a service is provided by one economic unit for another, no specific property right is actually exchanged between them in the way that the ownership of goods is transferred from one unit to another. Since models of

pure exchange economies assume property rights exist, the application of these models to service industries deserves careful scrutiny.

It has been argued by Carroll, Ciscil, and Chisholm [7] that there is a paradox in the private enterprise system: that while property rights are an essential precondition for the existence of a market, in many contexts property rights and market efficiency are tradeoffs. This is because competitive markets function efficiently when new entrants increase market output, forcing established firms to cut their price, their volume, and their margin of profit. Since the price-reducing consequence of market entry is shared by all sellers, new firms take no account of how their entry impacts on existing firms. This efficient functioning of the market thus generates negative externalities which is paradoxical: the essence of property rights is the ability to exclude while the chief characteristic of an ideal market is that no one can be excluded.

The literature on rent-seeking is extensive as shown by Hartle [12] or Tollison[16]. The literature on preclusive competition as it relates to common property resources is also extensive as seen, for example, in DeAlessi [10]. However, very little work has been done in terms of preclusive competition as it relates to markets which do not involve unappropriated natural resources. Copes [8] attempted to demonstrate such preclusive competition in an analysis which drew a parallel between the existence of excess capacity in certain markets and the occurrence of excess inputs in fishing industries operating under open access conditions. Drawing on the work of Carroll et al., he assumed the market in question, as an institution, was a commons and concluded that any open access market would result in a nonoptimal volume of both capacity and output and that there would be efficiency gains for society if entry were restricted. This is a significant break with traditional economic theory which assumes the existence of smoothly functioning markets. Traditional theory assumes the absence of transactions costs in the market regardless of their form.

Copes endeavoured to introduce transactions costs into the market in the form of adverse interactions among selling firms. We have demonstrated [5],that the Copes open access model is applicable if there is simultaneity of production and consumption of output. We made particular reference to the service sector because in service sector markets, customers are integral inputs into the production activity of the firm in that they are required in order for sales (output) of the firm to occur. We have also shown [6] that the finite availability of potential passengers creates problems when we adapted the Copes open access model to the analysis of a deregulated airline industry. We showed that open access leads to the dissipation of market rents.

3.2 Implications for Urban Passenger Markets:

Service sector firms compete with one another for customers in any market. The number of such customers (and their expenditures) is finite, though its availability can be influenced to some extent by price competition, terms of sale, or service quality. The objective is to attract customers either from rivals, or

into the market from other activities. Viewed from this perspective, the customer is a prize. A difficulty emerges: attempts to attract customers create adverse interactions among competing firms in a given market. The adverse interactions arise from the fact that the customers are available to whichever firm is best able to attract them. Once attracted by this firm, they are no longer available to another competitor.

The demand by passengers for urban transportation services can be considered as either a derived demand or a case of joint demand. In the case of the former, the need for transportation derives from some previous decision concerning work/home location or some consumption activity such as recreation or shopping. The transportation decision is secondary since the decision to travel has already been made. The interesting questions are those of modal split, and how variations in fares and/or quality of service can influence it. In the latter case, only consumption activities are relevant, and the two decisions are arrived at simultaneously. One still faces the problem of complementarity in that one of the goods tends to dominate. As before, transportation is a secondary issue, less important that the other. It is thus possible to argue that the total number of trips taken is determined exogenously. The total number of trips taken within a given time period is thus finite, and the problem of market congestion is thus relevant.

This is not to say that transportation decisions are not important. At least two dimensions need to be incorporated into the analysis. Trip destination, at least for one direction when consumption activities are involved, is a question which needs to be addressed. Such would not be the case for the journey to work unless one incorporated intermediate stops. The real issue, however, is that of modal split. That such applies to the journey to work is clear.

However, it is to be noted that it applies particularly to all recreational travel, even the annual vacation. Unlike other situations of joint demand, the two decisions will be reached sequentially rather than simultaneously. From our perspective, the dominant dimension here is the destination. Consider the annual vacation as a case in point. The choice of destination may well circumscribe the modal choice decision. Travel distance may limit choice in such a way as to preclude the automobile.

So it is with urban transportation. The location of home, work, or travel destination may well circumscribe modal choice decisions. In those cases where an individual has access to both automobile and bus transport, the disaggregate choice models may be used to assist in the estimation of the demand for each mode. Both price and quality of service dimensions can be incorporated. Other dimensions, such as time-of-day travel must be regarded as largely predetermined. What is important to note here is that, since the demand for transportation is a derived demand, there will be no change in the total number of trips within the urban environment in the short run. In a longer run context, one could expect the total amount of travel to vary directly with population (and associated age distribution), housing location, variations in real all-inclusive transportation costs and changes in real incomes.

This fixity of total passenger volumes leads to market congestion, and market congestion reduces contestability. Among specific routes in a system, some will be more contestable than others. At this level, we are really dealing with the contestability of particular markets, and both size and density are relevant dimensions. Such contestability depends upon the ability to sub-divide the network into components. It must be recognised that some portions of the urban market are too small to be considered contestable. Light density routes are not profitable, and are unlikely to be so.

4.0 THE MODEL

In the model, capacity is offered in order to service potential customers. With deregulation, access is open, and hence additional capacity may be added. If there is more than one firm, head-to-head (or service) scheduling develops as rivals attempt to maximise their passenger volumes. Incumbents may also add capacity so as to discourage entry. The result is an excessive amount of capacity at least in the more profitable segments. As capacity is added, on any given route, customer density declines until average returns drop to the level where there are no net benefits over and above normal returns. The addition of further capacity is no longer attractive.

The production relationship defines the number of passenger trips completed per period in a given market, at different levels of capacity, measured by the number of flights operated. The shape of the production function indicates that when access is open capacity will be added as long as competitors perceive available passengers. The limit on passenger availability implies that when more capacity is added, more trips are taken but at a declining marginal rate of increase as unserved passengers become progressively more scarce. Thus, both the average and marginal yield curves (in terms of passenger trips) are negatively sloped for the industry. Assuming a constant average price for passenger trips, the production curve may be transformed into a total revenue curve. The marginal and average curves can then be viewed in dollar terms.

For a fixed fleet, there are rising opportunity costs because additional capacity may be drawn into the market at rising cost or because the industry must utilise capacity units of lesser and lesser efficiency. The supply curve of capacity is upward-sloping and represents the marginal opportunity cost of capacity curve. On the other hand, total pecuniary costs are defined as those costs, both explicit and implicit, which producers must pay for the services necessary to produce a specified amount of capacity. Total opportunity costs and total pecuniary costs are equal if the marginal opportunity cost of capacity is constant. Otherwise, the difference between TOC and TPC represents the intramarginal rents earned by the more efficient capacity.

Implicit costs are all opportunity costs to the carriers from operations on the route. For a fixed fleet, these opportunity costs increase with capacity. The result is a positively sloped average pecuniary cost curve for the industry. The

average pecuniary cost is equal to the marginal opportunity cost which means an upward sloping supply of capacity (marginal opportunity cost) curve for the industry. The average opportunity cost curve lies under the average pecuniary cost curve, the difference between the two being the intramarginal rent.

Due to the open access nature of the industry, competitors equate the value of the average yield with average pecuniary cost. This is market rent dissipating rather than maximising behaviour. Market rents are dissipated because individual firms do not consider the effect that additional capacity will have on the productivity of the rest of the industry. However, intramarginal rents accrue to the more efficient operators in the industry. The result of the market rent dissipation is more capacity than is economically efficient. The rent maximising behaviour in the industry would see carriers add additional capacity only as long as its marginal value exceeded the marginal opportunity cost. Competitors would then maximise both market and intramarginal rents. This maximising behaviour is characterised by exclusive rights rather than open access. These effects are illustrated in **Figure 1. below**

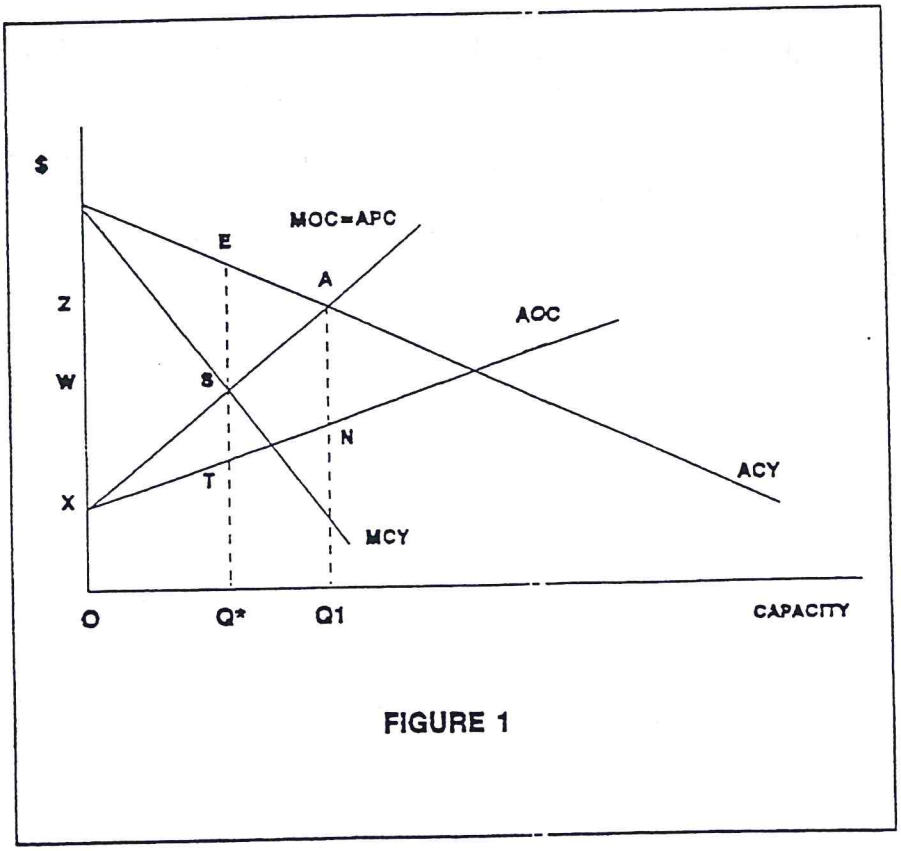

FIGURE 1

The downward sloping value of marginal capacity yield and value of average capacity yield curves at the industry level are labelled MCY and ACY respectively. MOC refers to the marginal opportunity cost of capacity in the industry and is equivalent to average pecuniary cost (APC). Optimal capacity is OQ* where marginal opportunity cost is equated with the value of marginal capacity yield. Intramarginal rents of ST x OQ* (or XSW) and market rents of ES x OQ* accrue to the industry. The market rent is attributable to the existence of market institutions (resulting from the removal of the regulatory constraints) while the intramarginal rent is due to operator efficiency. Under open access, equilibrium capacity will be OQ1 and all market rent is dissipated. This represents inefficient exploitation of the market. However, intramarginal rents of AN x OQ1 (or XAZ) now accrue. As suggested by Copes [10], the open access to the market has led to excess capacity. This excess capacity is a form of technological externality.

5.0 CONCLUSIONS:

This paper has considered the question of the contestability of urban passenger markets. It was determined that the problems confronting a deregulated airline industry in the US would not be likely to preclude contestability in deregulated surface transportation. The major problem here, and in retrospect for the other case as well, is what might be termed market congestion. Any open access scenario where the service sector is involved will result in over exploitation of the market(s) with resultant excess capacity. We show that this outcome is inevitable, and it is this outcome which reduces the contestability of such markets.

REFERENCES

[1] Bailey, Elizabeth E. and Jeffrey R. Williams. "Sources of Economic Rent in the Deregulated Airline Industry." Journal of Law and Economics, XXXI (1988), pp. 173-202.

[2] Baumol, W. J. Panzar and R. Willig, Contestable Markets and the Theory of Industry Structure, (New York, Harcourt, Brace Jovanovich, 1982).

[3] Baumol, W. and R.D. Willig, "Contestability: Developments Since the Book", Oxford Economic Papers, Vol. 38(1986): 9-36.

[4] Brander, J.R.G. and B.A. Cook. "Air Transport Deregulation and Airport Congestion: The Search for Efficient Solutions."Transportation Research Record, 1094 (1987), pp. 18-24.

[5] Brander, J.R.G. and B.A. Cook. "The Market as a Commons: A Further Comment." De Economist, 134 (1986), pp. 81-92.

[6] Brander, J.R.G., B.A. Cook and J.E. Rowcroft. 1989. "Deregulated Airline Markets as Open Access Commons." De Economist 137.

[7] Carroll, T.M., D.H. Ciscil and R.K. Chisholm. "The Market as a Commons: An Unconventional View of Property Rights." Journal of Economic Issues, XIII (1979), pp. 605-627.

[8] Copes, P. "The Market as a Commons: A Neglected Aspect of Excess Capacity." De Economist, 132 (1984), pp. 49-60.

[9] Coursey, D. et al., "Market Contestability in the Presence of Sunk (Entry) Costs", Rand Journal of Economics, Vol. 15, 1984.

[10] De Alessi, Louis. 1980. "The Economics of Property Rights: A Review of the Evidence." Research in Law and Economics 2 .

[11] Eads, G.C., "Airline Competitive Conduct", Antitrust Bulletin, 1983.

[12] Hartle, D.G. 1983. "The Theory of Rent Seeking - Some Reflections." Canadian Journal of Economics (Nov.) XVI.

[13] Meyer, J.R. and W.B. Tye, "Toward Achieving Workable Competition in Industries Undergoing a Transition to Deregulation: A Contractual Equilibrium Approach", Yale Journal on Regulation, 1988.

[14] Pigou, A.C. Economics of Welfare, first ed., 1920.

[15] Shepherd, W.G., "Contestability vs. Competition", AER, 1984.

[16] Tollison, Robert D. 1982. "Rent-Seeking: A Survey." Kyklos.

[17] Tullock, Gordon. 1967. The Welfare Cost of Tariffs, Monopolies and Theft." Western Economic Journal (June).