17

Journal of Transport Economics and Policy, Volume 39, Part 1, January 2005, pp. 25–52

Calculation of Economies of Spatial Scope from Transport Cost Functions with Aggregate Output with an Application to the Airline Industry

Leonardo J. Basso and Sergio R. Jara-Díaz

Address for correspondence: Professor Sergio R. Jara-Díaz, Department of Civil Engineering, Universidad de Chile, Casilla 228-3, Santiago, Chile. Leonardo Basso is also a faculty member at the University of Chile. This research was partially funded by Fondecyt, Chile, Project 1010687 and the Millennium Nucleus "Complex Engineering Systems". The authors wish to thank Cristián Cortés, Gabriela Morales, and three anonymous referees for helpful comments.

Abstract

Transport network expansions have usually been analysed calculating returns to scale with variable network size (*RTS*), which has been shown to suffer from a number of shortcomings because, in the end, it attempts to capture as a *scale* property something that in fact is related with *scope*, namely the addition of new products when a transport network expands. In this paper the authors develop a method to calculate economies of spatial scope from transport cost functions with aggregate output, which is then illustrated using the results of a published study on airlines. We conclude that the method holds very well and that, coupled with the strict calculation of economies of scale (corrected returns to density), it permits a clear explanation of observed firm behaviour.

Date of receipt of final manuscript: July 2004

1.0 Introduction

When a network variable *N* is included in the estimation of a transport cost function, empirical studies of the transport industry structure make a distinction between returns to density (*RTD*) and returns to scale (*RTS*). The former assumes a constant network when output increases (increase in density), while the latter assumes that the network grows as well (increase in output through a network expansion but keeping density constant). Accordingly, *RTD* is calculated as the inverse of the sum of product elasticities of cost, while *RTS* includes the network elasticity in the summation as well. Most empirical studies of the airline industry (where the number of points served, *PS*, is the usual network variable), have reported the presence of increasing returns to density and constant returns to scale, as concluded by Caves, Christensen, and Tretheway (1984), Kirby (1986), Gillen, Oum, and Tretheway (1985, 1990), Oum and Zhang (1991), Kumbhakar (1992), Keeler and Formby (1994) and Baltagi, Griffin, and Rich (1995) among others. These results indicate that, on costs grounds, it would be advantageous for firms to increase traffic densities on their networks, but it would be inconvenient to expand their networks. Observed industry behaviour, however, was different: after deregulation — in the US first and then in the rest of the world — the air industry has concentrated and the networks served have expanded through mergers, alliances, and acquisitions. Thus, firms have tried to increase their network size, which seems to contradict constant returns to scale as previously defined. Two reactions emerged; on the one hand, some authors have argued that network growth can be understood as an attempt to exploit economies of traffic density (for example: Oum and Tretheway, 1990; Brueckner and Spiller, 1994). On the other hand, a series of re-examinations of the methods to calculate scale economies for all transport industries have been proposed in the literature (Gagné, 1990; Ying 1992; Xu *et al.*, 1994; Oum and Zhang, 1997).

The re-examination of the calculation of scale economies has been mostly based upon the interrelation among the arguments of the estimated cost functions, namely products, attributes and network variables. This approach, however, has suffered from an important difficulty: no single set of output descriptions has been used, which is reflected in the lack of accepted standard definitions for economies of density and economies of scale. As some have pointed out, what is scale in one study may be density in another, and some have pointed out that "economies of scale, as they are usually measured in the context of manufacturing industries, simply represent economies of density in a network industry" (Filippini and Maggi, 1992, p.308). Jara-Díaz and Cortés (1996, from now on JDC)

proposed a new look at the subject showing that, in a rigorous sense, an improved version of what today is understood as economies of density is in fact scale under a strict definition, which had been suggested by Panzar (1989) using a linear aggregate output. The JDC approach was based upon the interpretation of the different forms in which output has been described as implicit representations of the displacements of goods and persons that transport firms produce. Let us briefly review this.

In the case of transport, a firm produces movements of different things between many origins and destinations (OD pairs), during different periods. Strictly, then, the product of a transport firm is a vector $Y = \{y_{ijkt}\}$, where y_{ijkt} represents flow of type k (goods or persons), between origin i and destination j, during period t (Jara-Díaz, 1982a,b; Winston, 1985; Braeutigam, 1999). JDC and many others have noted that as vector Y can not be used in the empirical work because of its dimension, estimated cost functions are specified in terms of aggregates that represent both products — such as ton-kilometres, total passengers, seat-miles — and so-called attributes — such as average length of haul or average load factor. Although both the strict description of transport output and the need to use aggregates for econometric purposes have been frequently recognised in the literature,[1] JDC stressed a fundamental fact, namely that behind these aggregates hides vector Y as defined above, and that this can and must be recognised when undertaking economic analyses. As economies of scale analyse cost behaviour as all outputs expand by the same proportion (Panzar and Willig, 1977), a correct calculation of scale economies in transport activities should be related with the same growth of all flows in vector Y (something first hinted by Griliches, 1972). But this could be analysed from an estimated cost function $\tilde{C}(\tilde{Y})$ — where $\tilde{Y} = \{\tilde{y}_1, \ldots, \tilde{y}_h, \ldots, \tilde{y}_{H^0}\}$ is the vector of aggregates — by examining the behaviour of each aggregate $\tilde{y}_h$ when Y varies. If aggregates are formally written as functions of Y, that is, $\tilde{y}_h \equiv \tilde{y}_h(Y)$, then $\tilde{C}(\tilde{Y}) \equiv \tilde{C}(\tilde{Y}(Y))$. From this basic observation, and by looking at scale in terms of the elasticities of cost with respect to the elements of Y, JDC showed that an internally consistent calculation of the degree of scale economies can be obtained using the elasticities of cost with respect to the aggregates $\tilde{y}_h$, provided each of these elasticities is weighted by its

[1] For example, Gillen *et al.* (1990, p.13) state that "ideally one would treat every type of airline service in every city pair market as a separate output. The impossibility of estimating a cost function with thousands of outputs requires some aggregation of the data. After aggregation, output attribute variables are introduced to control for some of the aggregation bias." Braeutigam (1999, p.68) states that "treating the movement of each commodity from each origin to each destination as a separate product would be desirable. There would be so many outputs however, that estimating a cost function would be impossible." See also Winston (1985, p.64–65) and Small (1992, p.50).

(local) degree of homogeneity with respect to Y.[2] In what follows, we will call vector Y the *true product* following Jara-Díaz (1982a) and Winston (1985), to distinguish it from the aggregate product used in empirical work.

Oum and Zhang (1997) pointed out that the JDC method for the calculation of economies of scale corresponds to an improvement of what is called *economies of density*, *RTD*, in the transport economics literature, because network variations (changes in N) are not allowed.[3] They are right: scale in terms of the true product Y — from now on referred as *disaggregated scale* — is in fact *density* in terms of $\tilde{Y}$. What about *RTS*, the relevant index when analysing network growth? In this direction, the JDC formulation goes beyond the correct calculation of returns to density/disaggregated scale. By looking at $\tilde{Y}$ in terms of Y, Basso and Jara-Díaz (2002) have shown that *economies of scale with variable network size*, *RTS* (*scale* in terms of $\tilde{Y}$ and N) are inadequate to study the cost effects of network expansions. In essence, this happens because *RTS* imposes that the traffic density — roughly the average load per link — remains constant after the network expansion, a condition that looks reasonable when considering $\tilde{Y}$ but is shown to impose unreasonable relations among flows when considering Y. Moreover, and in spite of proposals to depart from the constant density assumption, Basso and Jara-Díaz also show that improving *RTS* is a hopeless course of work.

As stated above, some have argued that the network growth observed in spite of $RTS = 1$ can be understood as an attempt to exploit economies of traffic density. For example, Brueckner and Spiller (1994) state that "the growth of networks can be understood as an attempt to exploit economies of traffic density, under which the marginal cost of carrying an extra passenger on a nonstop route falls as traffic on the route rises. By funneling passengers through a hub airport, the switch to hub and spoke operations raised traffic densities and allowed carriers to reduce their costs." Oum and Tretheway (1990) provide a graphical example to support their claim that "in spite of constant returns to scale/network size, the addition of a station to a hub and spoke system can result in economies. This is due to the exploitation of economies of traffic density by the traffic stimulation effects of hubs." Although in principle the argument seems reasonable, the increasing returns to density found in many studies were calculated explicitly keeping the size of the network fixed. This means that economies of density can be used without ambiguity to explain the merging of firms

[2]These "weights" happen to be between 0 and 1 for usual output aggregates.

[3]The theoretical reason for this is that an equiproportional change in Y produces no change in N. In other words, an increase in all OD flows provoke neither increase nor decrease in the number of points served, *PS*.

that serve the same set of nodes but, as found in every econometric study that has considered a network variable, expanding the network is costly[4] and this is not considered in the density justification for network growth. In our opinion, perhaps this is an intuitively correct but incomplete explanation: it does not take into account the necessary change in costs provoked by a larger network and there is no methodology, to our knowledge, to test the claim empirically. Note that *RTS* in its usual form might capture the fact that increasing the network is costly, but it cannot encompass the density explanation, since density is kept constant by construction.

What is next, then? If the concept of *economies of scale with variable network size*, *RTS*, is unsuitable to do a proper analysis of the cost effects of network growth and the *density explanation* cannot be tested and actually leaves outside the network expansion effect itself, what is to be done in order to examine the very relevant issue of network expansion in the transport industries? In our opinion, the situation so far is the following: we know how to detect the presence of economies of density/disaggregate scale, but we do not have a method to analyse the cost convenience of network expansions (or reductions) from cost functions estimated in terms of aggregated products. A new method, that replaces *RTS* in this task, is needed. We believe that the correct approach is the calculation of economies of spatial scope, because "increasing network size is unambiguously associated with an increase in the number of products and, therefore, networks variables are related with economies of scope" (Jara-Díaz, Cortés, and Ponce, 2001). This is particularly evident when N is represented by the number of points served, *PS*, because increasing *PS* implies increasing the number of OD flows. Economies of spatial scope have been suggested in the literature as an explanation for merging or the formation of alliances (see, for example, Hurdle *et al.*, 1989; Oum, Park, and Zhang, 2000). Nevertheless, although theoretical analyses about how and why economies of spatial scope could arise are available (see, for example, Jara-Díaz and Basso, 2003), there is no method to calculate them from an estimated cost function with aggregate output, to our knowledge. Note also that learning the extent of economies of scope is of interest not only to assess the cost convenience of network expansions, but also because economies of scope and scale are related with subadditivity, that is, with the existence of natural monopoly.

In this article we present a method to calculate spatial scope from transport cost functions with aggregate output, including an empirical example. The method also rests on the $\tilde{y}_h \equiv \tilde{y}_h(Y)$ property of aggregates,

[4]For example, Gillen *et al.* (1990) found an elasticity of the cost with respect to *PS* of 0.132.

proposed by JDC. But, just as behind *RTS* lies a constant density imposition in order to control for the value of aggregated flows after a network expansion, our method to calculate scope will impose a related condition on the disaggregated flows. The method not only permits us to study the cost convenience (or inconvenience) of network growth, but also provides a way of incorporating the density explanation in a strict economic way by considering cost changes produced by both density and network size increases. In order to illustrate the applicability and potential of the method, it is applied using information from a particularly rigorous published article on airlines. Using that data and results, we show how the analysis of industry structure can be improved and what kind of insights can be gained.

2.0 The Method

2.1 Economies of spatial scope

In multiproduct theory (Baumol, Panzar, and Willig, 1982), the degree of economies of scale S deals with the effect on costs of a proportional expansion of all products. On the other hand, economies of scope are related with the cost convenience (or inconvenience) of jointly producing two sets of products. If $C(Y)$ represents a cost function where input prices have been omitted for simplicity, the degree of economies of scope, SC, is calculated as

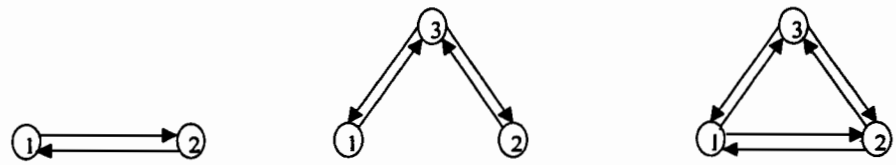
$$SC^A = SC^B = \frac{C(Y^A) + C(Y^B) - C(Y^D)}{C(Y^D)}, \tag{1}$$

where D is the set of all products, $A \cup B = D$ and $A \cap B = \varnothing$ (that is, A and B form an orthogonal partition of D). Y^A is vector Y^D with $y_i = 0$, $\forall i \notin A \subset D$; Y^B is defined analogously. Thus, a negative value for SC^A implies that it is cheaper to let a second firm produce the outputs in B than to expand the line of production of a firm already producing the outputs in A. If SC^A is positive, then it is cheaper to let a single firm produce everything (Y^D). Note that SC lies, theoretically, in the interval $[-1, 1]$, since it represents the proportion of cost savings due to joint production.

Emphasising the spatial dimension of product — or equivalently, assuming one type of cargo and one period for simplicity — in the case of transport S takes care of cost behaviour as OD flows expand proportionally, keeping the number of OD pairs constant, while SC deals with costs when OD pairs are added. Thus, within the context of variable networks, *spatial scope* is useful to analyse whether a certain firm A serving PS^A

Figure 1
Variable Network Size and Spatial Scope

$[A]$ $Y^A = \{y_{12}; y_{21}; 0; 0; 0; 0\}$ $[B]$ $Y^B = \{0; 0; y_{13}; y_{31}; y_{23}; y_{32}\}$ $[D]$ $Y^D = \{y_{12}; y_{21}; y_{13}; y_{31}; y_{23}; y_{32}\}$



nodes with $PS^A \cdot (PS^A - 1)$ potential OD flows should expand its network to serve PS^D nodes, adding $PS^D \cdot (PS^D - 1) - PS^A \cdot (PS^A - 1)$ new flows, or these new flows should be served by a new firm. To be precise through an example see Figure 1, where firm A serves two OD pairs and SC can be used to analyse the convenience of adding four new OD pairs through the addition of a new node (3). A positive value for SC would indicate incentives to add node 3, producing vector Y^D. Note that the incremental cost of serving a new node is given by $C(Y^D) - C(Y^A)$, which will be different from $C(Y^B)$ unless SC is nil.[5] Note also that the underlying physical network (links and terminals) that may be necessary to produce the flows is not shown. Furthermore, nothing is imposed about the route structure, that is, the way in which OD flows are produced using the network, for example directly, cyclically, or through a hub. What is shown is simply the OD structure of production. The role of the physical network as an *input*, of the route structure as an endogenous decision and the technical analysis behind transport operations that leads to cost functions, are presented in Jara-Díaz and Basso (2003).

2.2 Scope from aggregates: the role of the average OD flow

For the empirical estimation of transport cost functions, the use of aggregate descriptions of transport product is, in almost all cases, unavoidable. This implies that the magnitude of flows at an OD level is lost. Unless in the econometric work one keeps track of origin and destinations by using spatially disaggregated flows,[6] the calculation of SC directly from equation (1) is not feasible. However, most aggregates $\tilde{y}_h$ are implicit functions of Y. For example, if ton-kilometres (TK) and average length of haul (ALH) are

[5]For the n-nodes case see Jara-Díaz *et al.* (2001).
[6]See for example Jara-Díaz (1988) and Jara-Díaz, Donoso, and Araneda (1991).

considered in $\tilde{Y}$, then

$$TK(Y) = \sum_{ij} y_{ij} \cdot d_{ij} \tag{2}$$

and

$$ALH(Y) = \sum_{ij} y_{ij} \cdot d_{ij} \Big/ \sum_{ij} y_{ij}, \tag{3}$$

where d_{ij} is the distance travelled by flow y_{ij} between origin i and destination j (JDC analyse many other aggregates). Therefore, even if the true (disaggregated) product vectors Y^A, Y^B, and Y^D were unknown, SC could still be calculated correctly if the corresponding aggregates $\tilde{Y}(Y^A)$, $\tilde{Y}(Y^B)$, and $\tilde{Y}(Y^D)$ were known, and an estimated cost function $\tilde{C}(\tilde{Y}, PS)$ was available (Jara-Díaz, Cortés, and Ponce, 2001). Analytically, SC could be calculated through

$$SC^A = SC^B = \frac{\tilde{C}(\tilde{Y}(Y^A), PS^A) + \tilde{C}(\tilde{Y}(Y^B), PS^B) - \tilde{C}(\tilde{Y}(Y^D), PS^D)}{\tilde{C}(\tilde{Y}(Y^D), PS^D)} \tag{4}$$

where $PS^D = PS^B > PS^A$ (see Figure 1). Note that the arguments of $\tilde{C}(\tilde{Y}, PS)$ in equation (4) are not evaluated at zero output levels, unlike $C(\cdot)$ in equation (1), allowing the use of translog cost functions directly for the calculation of scope.[7] This happens because the aggregate representations (such as total passengers or ton-kilometres) do not vanish when some of the OD flows go to zero as in Y^A or Y^B. Furthermore, note that although Y^B and Y^D are associated with the same number of points served, the number of OD flows is different, which makes the aggregate descriptions different, that is, $\tilde{Y}(Y^B) \neq \tilde{Y}(Y^D)$ and $\tilde{C}(\tilde{Y}(Y^B), PS^B) < \tilde{C}(\tilde{Y}(Y^D), PS^D)$ (Jara-Díaz, Cortés, and Ponce, 2001).

Let us explore how can we use equation (4) to calculate economies of spatial scope from an estimated aggregated cost function. As explicitly considered in equation (4), behind the vector of aggregates $\tilde{Y}^A$ there is a true flow vector $Y^A = \{y_{ij}^A\}$ that generates the aggregates. Although the values of the y_{ij}^A flows remain hidden behind the (known) aggregates, we know they fulfil equations such as (2) and/or (3). PS^D is also known since we are analysing a network expansion from PS^A to PS^D. By construction, PS^B is equal to PS^D. What remains unknown in equation (4) are $\tilde{Y}(Y^B)$ and $\tilde{Y}(Y^D)$. The former is the aggregated representation of the product vector implicitly containing the flows that are added after the network expansion. The latter is the aggregated representation of the

[7]Of course one has to maintain the assumption that the estimated cost function does a good job in describing costs, in spite of being specified with aggregate descriptions of product. In other words, $\tilde{C}(\tilde{Y}) \equiv \tilde{C}(\tilde{Y}(Y)) = \hat{C}(Y)$.

product vector implicitly containing both the original flows served — the y_{ij}^A flows — and the added flows.

In the absence of more information, some reasonable condition has to be imposed regarding the magnitude of the flows added after the network expansion, in order to assign values to the aggregates in both $\tilde{Y}^B$ and $\tilde{Y}^D$, given that one knows $(\tilde{Y}^A; PS^A)$. This type of requirement is not new; the condition behind the calculation of *RTS* that implicitly assigns values to $\tilde{Y}^B$ and $\tilde{Y}^D$ is that *density* does not change, where *density* is understood as the ratio between those aggregates whose elasticities were considered in the calculation, and the variable representing the network (*PS* in our case). As is evident, the condition behind *RTS* is imposed on the aggregates, not on the true flow vector Y. But what seems to be a reasonable condition on the aggregates $\tilde{Y}$ induces unreasonable analytical restrictions on the new OD flows (that is, those behind Y^B) that prevent a meaningful analysis of industry structure (Basso and Jara-Díaz, 2002). If a condition is needed in order to assess the impacts of a network expansion on costs, it should be imposed on the true product, the flow vector Y, even if one is working with aggregates for econometric purposes. This will permit consistent and more reasonable inferences later on. Therefore, we propose to calculate economies of spatial scope using equation (4), under the condition that the average OD flow of each cargo type remains constant after the network expansion. Formally, we define the average origin-destination flow for cargo type k as

$$AOD_k = \frac{\sum_i \sum_j y_{ijk}}{NOD}, \tag{5}$$

where y_{ijk} represents the flow of type k between origin i and destination j and *NOD* is the total number of OD pairs served. Note that the numerator in equation (5) is total tons T if k indicates freight and total passengers P if k refers to persons. This means that two indexes ought to be calculated if in the study in question freight and passenger services are provided. Holding this index (indexes) constant when calculating economies of spatial scope through equation (4) will permit the analytical estimation of the values of the components of both $\tilde{Y}^B$ and $\tilde{Y}^D$. The idea is simple: calculate AOD_k^A from $(\tilde{Y}^A; PS^A)$, and then estimate $\tilde{Y}^D$ and $\tilde{Y}^B$ from $AOD_k^D \equiv AOD_k^A$.

This general approach will be procedurally dependent on the aggregates used in each application; in what follows we show analytically how the method works for a particular although very common aggregate cost function and then we present an application. It is important, however, to stress at this point what the proposed approach accomplishes conceptually.

First, it should be clear that RTS is being replaced by something more adequate. On one hand, the condition we propose is explicit and made in terms of the true product Y, as opposed to the constant density assumption. On the other hand, we are using the concept specifically built to study the cost convenience of adding new products: economies of scope. Second, the method can accommodate the intuitive density explanation for network growth, but taking into account the costs associated to the larger network and providing an empirical way to test it. To see this, note that some of the new OD flows — which will be on average as large as the originals — will circulate in the original portion of the network, increasing density, provided they are not all directly served. This will be cost convenient if increasing returns to density are present. But, by calculating SC, the costs of the network expansion itself will be properly captured this time, as is evident from equation (4). It is not surprising that increasing returns to density favours the presence of economies of spatial scope even though they represent totally different ways of increasing output. After all, economies of density represent economies of scale in terms of the true product Y and it is well known that the presence of economies of scope favours economies of scale and vice versa because of a general theoretical property (Baumol, Panzar, and Willig, 1982). It must be clear, however, that even in the presence of decreasing or constant returns to density/disaggregated scale, economies of spatial scope may exist, as illustrated in Jara-Díaz and Basso (2003).[8] Finally, note that a proportional expansion of all flows in all OD pairs (which is what lies behind the strict notion of scale) makes AOD and density (as understood in the literature) grow by the same proportion. Therefore, the relation between density and disaggregated scale remains intact with the proposed approach.

2.3 Using AOD to calculate spatial scope from the aggregates: an example

How does this apply to an estimated cost function? Let us illustrate this general approach with a specific — although popular — example. Let us consider an aggregated cost function $\tilde{C}(TK, ALH, PS)$ where TK represents ton-kilometres and ALH represents average length of haul (defined in equations (2) and (3)). We would like to examine whether a certain firm A, with cost given by $\tilde{C}(TK^A, ALH^A, PS^A)$, has cost incentives

[8]This could happen because of the same alleged reasons why there might be increasing RTS, as for example shared use of airport and ground staff, handling of baggage transfers, and passengers check-in (see for example Oum, Park, and Zhang, 2000). Note, however, that $SC > 0$ and $RTD = 1$ cannot be parallelled by $RTS > 1$ and $RTD = 1$ since $RTS < RTD$ analytically (Basso and Jara-Díaz, 2003).

to connect new nodes (airports) assuming that the new OD flows have, on average, the same magnitudes as those already served (constant AOD), or if it is better to have another firm serving them. The expanded network will have a size given by PS^D, the incremental cost of serving the new flows is given by $\tilde{C}(TK^D, ALH^D, PS^D) - \tilde{C}(TK^A, ALH^A, PS^A)$, and the cost of producing these new flows with a different firm is $\tilde{C}(TK^B, ALH^B, PS^B)$. Then, replacing terms in (4), what should be calculated is

$$SC^A = \frac{\tilde{C}(TK^A, ALH^A, PS^A) + \tilde{C}(TK^B, ALH^B, PS^B) - \tilde{C}(TK^D, ALH^D, PS^D)}{\tilde{C}(TK^D, ALH^D, PS^D)}. \tag{6}$$

In (6), we only know TK^A, ALH^A, and PS^A, which represents the point where economies of spatial scope will be calculated. The value of PS^D will depend on the size of the network increase that we would like to study, for example one or five nodes. As RTS is a local marginal measure, we would like to consider a marginal change in network size, that is, one node, which implies $PS^D = PS^A + 1$. From this case, every network expansion can be analysed incrementally. We will assume that firms A and D potentially serve all their corresponding OD pairs, which means that

$$NOD^A = PS^A \cdot (PS^A - 1), \tag{7}$$

$$NOD^D = PS^D \cdot (PS^D - 1) = PS^A \cdot (PS^A + 1). \tag{8}$$

The average OD flow before the expansion is given by

$$AOD^A = \frac{T^A}{NOD^A}, \tag{9}$$

where T^A is total tons moved by firm A, that is, $T^A = \sum_{ij} y_{ij}^A$. As ALH is the ratio between TK and T (see equations (2) and (3)), equation (9) can be re-written as

$$AOD^A = \frac{TK^A}{ALH^A \cdot NOD^A}. \tag{10}$$

After the network expansion, the average OD flow is given by

$$AOD^D = \frac{TK^D}{ALH^D \cdot NOD^D}. \tag{11}$$

Now we are fully prepared to use the constant AOD condition — which plays a role that parallels constant density in RTS — to estimate the remaining unknown variables TK^D, ALD^D, TK^B, and ALH^B. Constant average OD flow implies equality between (10) and (11). Imposing this

and replacing equations (7) and (8) we obtain

$$TK^D = TK^A \cdot \frac{ALH^D}{ALH^A} \cdot \frac{PS^A + 1}{PS^A - 1}. \tag{12}$$

Equation (12) shows that assigning a value to TK^D requires an estimate of the (potential) variation of the average length of haul after the addition of one node to the network, ALH^D. Note that a constant average OD flow is a condition that implies nothing on travelled distances. This is not unexpected but consistent with a strict notion of output because, from a disaggregated viewpoint, the route structure, and therefore travelled distances and ALH, are the result of endogenous operational decisions that depend not only on the magnitude of all flows served but also on the network topology (Jara-Díaz and Basso, 2003). In fact, the literature on empirical transport cost functions offers explicit alternatives to deal with ALH when the network expands. One alternative is to consider it constant, a choice implicitly or explicitly made in many articles where economies of scale with variable network size have been calculated (see, for example, Kumbhakar, 1992). A second alternative is the empirical examination of the variation of ALH with the network variable, following the proposition of Oum and Zhang (1997). In this alternative, if data permits, the idea is to estimate a function f such that $ALH = f(PS)$. Obviously, both alternatives for ALH^D can be explored.

Next, we need estimates for TK^B and ALH^B. The former could be estimated as

$$TK^B = TK^D - TK^A, \tag{13}$$

that comes from

$$TK^B = \sum_{ij} y_{ij}^B \cdot d_{ij} = \sum_{ij} y_{ij}^D \cdot d_{ij} - \sum_{ij} y_{ij}^A \cdot d_{ij} = TK^D - TK^A, \tag{14}$$

which implies that the distances travelled by the original flows (served by A) do not vary after the new node is added. As the route structure is an endogenous firm decision, distances could change if the route structure is changed. Two things should be noted. First, assuming the distances travelled by the original flows do not change ensures that, after the network expansion, the incorporation of the new flows will increase the density on the original portion of the network, as long as the new OD pairs are not served exclusively by direct services. This phenomenon can be easily portrayed using a hub-and-spoke route structure, but certainly is not limited to that case. Second, ALH may still change because the new flows do not need to travel, on average, the same distances as the original flows. On the other hand, the equality $T^B = T^D - T^A$ stands without discussion.

Once TK^B has been calculated, the respective average length of haul can be obtained from

$$ALH^B = \frac{TK^B}{T^B} = \frac{TK^B}{T^D - T^A} \Rightarrow ALH^B = \frac{TK^B}{\dfrac{TK^D}{ALH^D} - \dfrac{TK^A}{ALH^A}}. \qquad (15)$$

Note that equation (15) for ALH^B is valid in general, irrespective of the alternative chosen for ALH^D. In particular, if $ALH^D = ALH^A$, then $ALH^B = ALH^A$.

For synthesis, the approach to calculate scope applied to the cost function $\tilde{C}(TK, ALH, PS)$ in our example turns into the following sequence of calculations

1. $PS^D = PS^A + 1$
2. $PS^B = PS^A + 1$
3. $ALH^D = f(PS^D)$ or $LH^D = ALH^A$
4. $TK^D = TK^A \cdot \dfrac{ALH^D}{ALH^A} \cdot \dfrac{PS^A + 1}{PS^A - 1}$
5. $TK^B = TK^D - TK^A$
6. $ALH^B = \dfrac{TK^B}{\dfrac{TK^D}{ALH^D} - \dfrac{TK^A}{ALH^A}}$
7. $SC^A = \dfrac{\tilde{C}(TK^A, ALH^A, PS^A) + \tilde{C}(TK^B, ALH^B, PS^B) - \tilde{C}(TK^D, ALH^D, PS^D)}{\tilde{C}(TK^D, ALH^D, PS^D)}.$

As indicated earlier, the sequence corresponds to a marginal increase of the network by one node, that is, from PS^A to $PS^A + 1$. This can be repeated to evaluate the evolution of SC as the network grows. As is evident, a different set of aggregated products in the cost function would require other sequence and specific calculations; for example, the use of passenger-kilometres mean that PK^D and PK^B should be calculated from PK^A taking into account average length of trip instead of average length of haul. Although a case by case analysis is necessary for other aggregates and attributes, the key aspect is to study analytically the behaviour of each one under the constant average OD flow condition.

The degree of economies of spatial scope calculated with the method proposed above has a main objective: to investigate whether there are cost advantages for the firms to expand their network size. This, in

conjunction with the degree of economies of density/disaggregated scale, will permit a correct analysis of the industry structure taking into account both density (level of production) and network size. In the following section the method will be applied using a published study on the Canadian air industry that reports an econometrically reliable cost function with aggregate output, which we found appropriate to explore the potential of our approach and its explanatory power.

3.0 Methodological Application

3.1 Summary of the selected article

In order to illustrate the methodological approach described in the previous section, we have selected the article by Gillen *et al.* (1990) — built on Gillen *et al.* (1985) — which was chosen for various reasons. First, one of the declared objectives of the authors is "to extend the knowledge on airline cost structure where there is significant variation in the size of carriers, and to clarify the issue of scale economies for small carriers." Second, the estimated cost function includes products and attributes that permit a quite direct application of the method presented. Third, the econometric work is carefully developed, which makes the quantitative results reliable for analysis. Finally, their calculation of *RTD* is a correct estimation of the (disaggregated) degree of scale economies, according to the Jara-Díaz and Cortés (1996) methodology. It is worth noting that this article is an extension of previous single output studies to an aggregated multiproduct case including freight, regular passenger services, and charter. In what follows we reproduce only those aspects that are necessary for our application and analysis.

Six regional and transcontinental Canadian airlines were observed annually from 1964 to 1980. Using a translog form and variables expressed as deviations from their means, a long-run cost function was specified as

$$\tilde{C} = \tilde{C}(F, PS, \phi(\tilde{Y}, Q), w, t), \tag{16}$$

where $\tilde{C}$ is total cost; F is a vector of *firm specific effects*; PS is the number of points served (network size); $\phi(\tilde{Y}, Q)$ is a vector of hedonic products; $\tilde{Y}$ is a vector of aggregate products; Q is a vector of attributes; w is a vector of input prices; t is time.

The aggregated product vector considered three components (our notation): scheduled revenue passenger-kilometres (PK), scheduled revenue freight ton-kilometres (TK), and non-scheduled (charter) revenue ton-kilometres including passenger and freight services (CH). Only the first

component was specified in a hedonic form (Spady and Friedlaender, 1978) as $\phi_1(PK, Q^1)$. Originally, the vector of attributes Q^1 considered was: average stage length for the scheduled passenger service (*ASL*), scheduled passenger load factor (*PLF*), and the weighted average number of seats per scheduled passenger service (*PNS*). The results showed that neither *PLF* nor *PNS* were statistically significant. Thus, the final long-run cost function included a three *outputs* vector (ϕ_1, *TK*, *CH*), with a log-linear specification for ϕ_1:

$$\ln \phi_1 = \ln PK + \beta \cdot \ln ASL. \tag{17}$$

Input prices were aggregated into three categories: labour, fuel, and capital and materials. Note that the cost function is estimated in terms of the average stage length (*ASL*) and not the average length of trip (*ALT*), which is what we used in the example in the previous section. *ASL* represents the average distance of a link in the route structure, where the average takes into consideration the number of passengers on each link rather than at an OD level.[9] For the application we gathered information on *ALT*, on the average length of trip for charter services (*ALT-CH*) and the average length of haul of freight (*ALH*). For the sake of simplicity, the method will be applied only to three of the six firms studied by Gillen *et al.* These carriers are Air Canada, the largest, and the relatively small Nordair and Quebecair. In Table 1 we show the 1980 values for their products, attributes, network, and input prices.[10] The mean values for 1980 and the overall means (deviation point) are also shown. The results reported by the authors for *RTD* and *RTS* in 1980 (calculated with *ASL* constant) are shown in Table 2.

From their results, the authors conclude that:

(1) Significant increasing returns to density (*RTD*) are observed at all data points. Unexploited economies of traffic density are larger for small carriers (Nordair and Quebecair) than for the larger carriers (Air Canada).

(2) Constant returns to firm network size (*RTS*) are found at all data points, except for Air Canada and Nordair, which exhibit decreasing and increasing returns respectively (which is the main reason why these particular firms were chosen for the example).

[9]Certainly, unless every OD pair is served directly, *PK* on links will not have a direct counterpart with *PK* on OD pairs. However, it is easy to check that *total PK* stemming from the addition of flows on links yields the same result as adding *PK* on OD pairs.

[10]In fact, these are input price indexes adjusted for a constant rate of technological change. See Gillen *et al.* (1990).

Table 1
Reported Airlines Production (aggregates) and Input Prices in 1980 and Overall Mean

	Air Canada	*Nordair*	*Quebecair*	*Mean 1980*	*Overall Mean*
PS	59	21	23	32.67	29.90
PK (millions)	23,767.71	703.61	289.70	6,028.17	3,143.30
ASL	1,114.60	528.37	273.49	682.03	590.19
ALT	1,832.89	734.51	402.1	1,127.85	
ϕ_1	3,781.31	136.12	66.60	1,090.76	590.73
CH (millions)	83.15	60.49	27.86	59.95	40.45
ALT-CH	3,606.50	2,345.01	1,454.91	2,661.59	1,560.80
TK (millions)	552.20	24.11	3.26	142.37	83.41
ALH	3,483.10	1,566.20	1,089.65	1,988.80	1,691.40
Labour price	1.93	1.80	1.69		
Fuel price	4.54	5.14	5.13		
Cap/Mat price	2.04	2.03	2.01		

Sources: Gillen *et al.* (1990), Gillen *et al.* (1985), and Statistics Canada *Transcontinental & Regional Air Carrier Operations* (years 1971–80), Catalog Number 51-001.

Table 2
Returns to Density and Returns to Scale in 1980

	Air Canada	*Nordair*	*Quebecair*	*Mean 1980*
RTD	1.147	1.263	1.209	1.211
RTS	0.881	1.147	0.993	0.971

Source: Gillen *et al.* (1990).

(3) The smaller carriers have a higher unit cost than the largest carrier, Air Canada, which is consistent with the observed consolidation of regional carriers into a single one.
(4) The finding of constant returns to firm/network size suggests that a small network carrier should not have a cost disadvantage, provided it achieves traffic densities within its small network similar to those of the larger carriers (Air Canada).

3.2 Application of the proposed method and industry structure conclusions

As three services are distinguished in this study, the sequence of calculations presented in Section 2.3 ought to be applied not only to *TK*, but also to *PK* and *CH*. Because of this, and to keep a clear picture of the firms we are about to analyse through economies of scope, we have calculated the average OD flow, *AOD*, for each firm and for each type of service in 1980. This is presented in Table 3.

Table 3
Average OD Flow (AOD) for Each Firm and Each Service in 1980

	Air Canada	*Nordair*	*Quebecair*
Passenger (pax)	3,789	2,281	1,424
Charter (tons)	6.74	61.42	37.84
Freight (tons)	46.33	36.65	5.92

It can be seen from Table 3 that charter service is more important than freight for Quebecair and Nordair, while for Air Canada the opposite applies, something pointed out by Gillen *et al.*[11] Regarding regular passenger transport, Air Canada presents an average flow that is one-and-a-half and two-and-a-half times that of Nordair and Quebecair respectively. In general, Air Canada presents larger OD flows and network size than Nordair and Quebecair (almost three times more points served than the others).

Although *ASL* is the only distance-related variable that needs to be inserted directly in the cost function (through ϕ_1 in equation (17)), the average length of haul of freight services (*ALH*) and the average length of trip for both passenger and charter services (*ALT* and *ALT-CH*) will be needed for other calculations, particularly when estimating the outputs of the *D*-like firms (see step 4 of the sequence). As explained in Section 2, average distances — *ASL*, *ALT*, *ALT-CH*, and *ALH* — can either be assumed to remain constant as the network grows or its variation with *PS* can be examined empirically, following Oum and Zhang's (1997) procedure, that is, log-linear equations of the form $\ln ASL = \alpha + \beta \ln PS$. Let us first use the second alternative. As detailed data are not known, approximated relations between *PS* and *ASL*, *ALT*, *ALT-CH*, and *ALH* can be guessed using the available information for the six firms considered in the study for years 1971, 1978, and 1980. From these 18 observations we obtain the results shown in Table 4 (standard errors in parentheses).

In order to obtain the average distances for the three *D*-like firms, the derivatives with respect to *PS* will be approximated as follows:

$$\ln ASL = \alpha + \beta \cdot \ln PS \Rightarrow ASL = \exp(\alpha) \cdot PS^{\beta};$$

$$\frac{\partial ASL}{\partial PS} = \exp(\alpha) \cdot \beta \cdot PS^{\beta - 1} \Rightarrow \frac{\Delta ASL}{\Delta PS} \approx \exp(\alpha) \cdot \beta \cdot (PS^{A})^{\beta - 1}.$$

[11] Gillen *et al.* (1985, p.27) state: "Regional carriers have operated substantial amounts of charter services. This is not surprising, given the lack of growth opportunities in the more highly regulated regional markets for scheduled traffic." Note that comparisons with passenger services are inappropriate since these are expressed in number of passengers while the other two are expressed in tons.

Table 4
Elasticities of Average Distances with respect to PS

	ASL	*ALT*	*ALT-CH*	*ALH*	*ASL*[a]
α	4.348	5.008	6.290	5.236	5.793
	(0.854)	(0.999)	(0.558)	(0.890)	(1.540)
β	0.598	0.534	0.505	0.644	0.461
	(0.258)	(0.302)	(0.169)	(0.269)	(0.101)
R^2	0.25	0.17	0.36	0.26	0.17

[a]Oum and Zhang (1997) results.

As ΔPS is equal to one, we finally have

$$ASL^D = ASL^A + \exp(\alpha) \cdot \beta \cdot (PS^A)^{\beta - 1}. \tag{18}$$

Similarly, three other equations are obtained for *ALT*, *ALT-CH*, and *ALH*, which will be used for the analogous step 3 of the sequence. Before calculating economies of spatial scope, though, there is a small issue to deal with. As can be seen in Section 2, ALT^B, $ALT\text{-}CH^B$, and ALH^B can be calculated on theoretical grounds (step 6 of the sequence). This is not true for *ASL*, however. To overcome this problem we impose that the ratio between ASL^B and ALT^B, is equal to the average of the ratios $ASL^A:ALT^A$ and $ASL^D:ALT^D$, based on the fact that this ratio is actually fairly stable in time — which is why the elasticities of *ASL* and *ALT* with respect to *PS* are alike. Sensitivity analyses are performed later.

Economies of spatial scope can finally be calculated. The coefficients of the calibrated translog cost function are reported in Gillen *et al.* (1990). The products for firms playing the role of *A*, *B*, and *D* for each of the three Canadian airlines and the value for *SC* are shown in Table 5. The results obtained for the degree of economies of scope are quite appealing. First, all three fall within the theoretical range, that is, between 1 and −1. Second, the results show that the three firms have increasing returns to spatial scope but smaller firms have stronger incentives to expand their networks considering constant average OD flows (*SC* being more than twice as big as Air Canada's). Recall that all three firms presented increasing *RTD* (see Table 2), which means economies of scale in the strict sense, that is, increasing production (OD flows) on a fixed network presents cost advantages. Third, the ordering of *SC* follows the same as *RTD*. According to the discussion in Section 2, by increasing their networks firms are able to partially exploit economies of density. The measure of *SC* however, does incorporate the fact that expanding the network is costly. Lastly, the ordering of *RTS* is also the same as *SC*, but only Nordair has *RTS* above 1. This may be seen as an indication that *RTS* could be

Table 5
Calculation of Economies of Spatial Scope (1980)

	Air Canada	*Nordair*	*Quebecair*
PS^A	59	21	23
PK^A	23,767.71	703.61	289.70
ASL^A	1,114.60	528.37	273.49
ALT^A	1,832.89	734.51	402.10
ϕ^A	3,781.31	136.12	66.60
CH^A	83.15	60.49	27.86
$ALT\text{-}CH^A$	3,606.50	2,345.01	1,454.91
TK^A	552.20	24.11	3.26
ALH^A	3,483.10	1,566.20	1,089.65
PS^D	60	22	24
PK^D	24,747.55	794.34	330.60
ASL^D	1,123.58	541.97	286.60
ALT^D	1,844.84	753.84	420.63
ϕ^D	3,928.93	152.65	75.08
CH^D	86.36	67.21	30.87
$ALT\text{-}CH^D$	3,620.69	2,368.68	1,477.54
TK^D	573.75	26.90	3.63
ALH^D	3,498.41	1,588.31	1,111.06
PS^B	60	22	24
PK^B	978.84	90.73	40.90
ASL^B	1,333.57	681.16	425.12
ALT^B	2,191.29	947.18	624.47
ϕ^B	148.73	16.42	8.38
CH^B	3.21	6.72	3.01
TK^B	21.55	2.79	0.37
SC_1	0.02811	0.08178	0.06247

biased *downwards* in terms of policy conclusions, that is, that it would show more often than SC that network expansions are not convenient.

Although these results provide a useful insight on the structure of the industry, one can well wonder whether factors other than network size have a stronger influence on them. The following three issues are explored, as a way to test the robustness of the results and to help us understand better what is behind the numbers.

(1) Airlines with smaller network sizes have higher returns to scope; however they also have higher returns to density/disaggregated scale, which favours economies of scope. This may partially explain the differences. In order to explore this issue further, we can compare the degree of economies of spatial scope of firms that have similar average OD flows, but different network size. The procedure to calculate these new values for SC is simple: we increase the values

of *PK*, *TK*, and *CH* for each firm, until they all reach the same average flow (hence given by the maximum).

(2) As explained earlier, the use of an econometrically obtained function to measure the change of the average distances, following Oum and Zhang (1997), is one of the alternatives for these variables. Alternatively, one can choose to impose that they remain constant (within each firm) when calculating spatial scope. This is particularly important to examine in this case because: (i) the regressions were performed with a rather limited data set; (ii) *ASL* was updated using a rule that may be reasonable but is not grounded on theory as the others are; and (iii) Gillen *et al.* calculated *RTS* assuming *ASL* remained constant. To calculate a new set of *SC* we simply impose that average distances do not change or, which is equivalent, that the elasticities are zero in Table 4.

(3) As can be seen from Table 1, the three airlines faced different input prices. While this certainly has an influence on the degree of economies of scope, we would like to know whether that influence is sizeable or not. To control for this factor we calculate new values for *SC* for each firm using the same set of input prices (the average).

These sensitivity analyses — and every possible combination of them — are presented in Table 6.

Regarding the first issue — the influence of output levels — it can be observed from measures 3, 4, 7, and 8, that even if similar production

Table 6
Economies of Spatial Scope (SC) in Various Cases

		Air Canada	*Nordair*	*Quebecair*
	SC with different input prices across firms			
SC_1	Average distances are variable	0.02811	0.08178	0.06247
SC_2	Average distances are fixed	0.02559	0.07635	0.05615
SC_3	Average distances are variable AOD are equal	0.02729	0.07130	0.07453
SC_4	Average distances are fixed AOD are equal	0.02486	0.06638	0.06660
	SC with identical input prices			
SC_5	Average distances are variable	0.02761	0.08203	0.06315
SC_6	Average distances are fixed	0.02513	0.07658	0.05673
SC_7	Average distances are variable AOD are equal	0.02680	0.07154	0.07525
SC_8	Average distances are fixed AOD are equal	0.02441	0.06659	0.06721

scales (in the true product sense) are imposed, firms with smaller networks still show higher returns to spatial scope (network size) than Air Canada. Comparing SC_3 to SC_1, and SC_7 to SC_5, reveals that the degree of economies of scope decreased for Air Canada and Nordair but increased for Quebecair. As explained before, increasing density (production) exploits part of the increasing returns to density (scale), pushing *SC* towards zero. Why then has Quebecair a higher measure of scope? The answer lies in the fact that *RTD* is a ray measure: it tells what happens when all the flows increase by the same proportion. In other words, the product mix remains unchanged. In this exercise however, this was not the case since, in order to impose equal *AOD* in every service, we necessarily needed to change output mix. For example, Quebecair's freight service was increased almost eight times while its passenger service was increased only two-and-a-half times. If we impose an equiproportional increase in all services equating passengers' average flow only, the degree of economies of scope indeed decreases for the three firms.[12] Firms with small networks still display higher returns to spatial scope in this application.

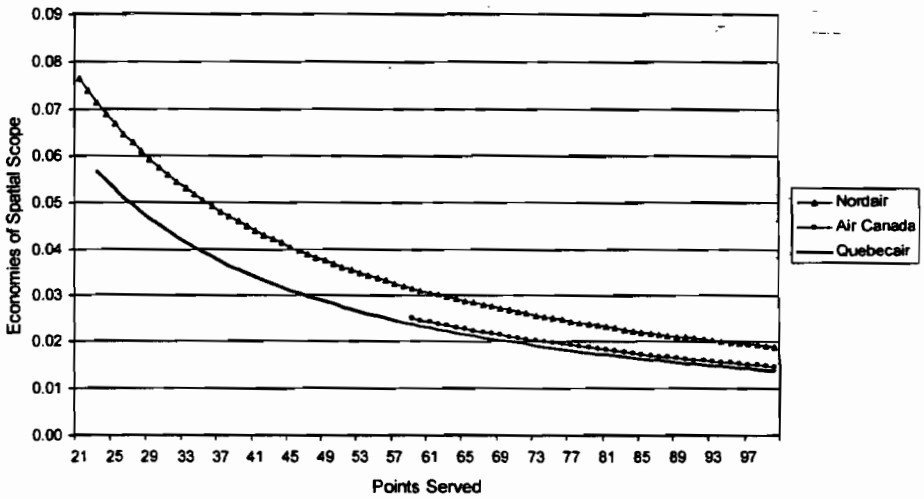
Regarding the second issue, it can be seen that when distances are imposed to remain unchanged within each firm (*SC* measures 2, 4, 6, and 8), the degree of economies of spatial scope decreases by around 9.8 per cent for Air Canada, around 7.4 per cent for Nordair, and around 11.5 per cent for Quebecair. This shows that conclusions remain basically unaffected. This is reassuring because, in the absence of detailed information, our econometric regressions for the average distance were somewhat unreliable. We know now that the potential impacts of this are small.[13] Finally, it can also be seen that the influence of input price differences is practically immaterial. The degree of economies of scope of Air Canada decreases by at most 1.8 per cent when its input prices are replaced by the average, while *SC* increases by 1 per cent for Quebecair and 0.3 per cent for Nordair.

In order to examine yet another angle of spatial scope, we have calculated *SC* using the proposed methodology increasing *PS* by one up to 100. We have done this considering constant average distances and identical input prices, to isolate better the effect of the network size in *SC* (that is, as in SC_6). The values obtained for each firm size are represented

[12] In this case, output levels of Nordair are increased 1.66 times and Quebecair's 2.66 times. For variable distances and different input prices, *SC* is 0.02779 for Air Canada, 0.07638 for Nordair, and 0.05302 for Quebecair.

[13] Moreover, our elasticities are larger than the elasticity suggested by Oum and Zhang (1997), which was calculated with detailed data (see Table 4). Our two analyses, then, can be considered as enclosing one carried out with that elasticity.

Figure 2
Economies of Spatial Scope as a Function of Network Size



in Figure 2. Recall that the values represented in this figure have been obtained by imposing that, for every service, average OD flows are constant as the number of points served increases.

Figure 2 suggests that the initially smaller returns to scope shown by Air Canada are due to its already large network. The other two firms can exploit their initially higher returns to spatial scope by increasing their network sizes. One would expect economies of spatial scope to diminish with network size, just as returns to scale diminish with production. Note, however, that this is a property of this cost function and its parameter estimates; it is not something that is guaranteed *a priori*. We therefore view the fact that returns to spatial scope diminish with *PS* in this case, as supporting both our method and the cost function estimated by Gillen *et al.* (1990).

The behaviour of *SC* for Nordair and Quebecair is quite similar to Air Canada's once they achieve large enough networks.[14] It is important to stress that this is so, in spite of the fact that they have quite different output mixes, average OD flows, and average distances. We can conclude then that these firms, which have smaller production scales and stage lengths, should not have disadvantages with respect to the larger ones as long as they reach sufficiently large networks; this is the main numerical

[14]The same pattern is observed when this incremental analysis is made either using the regressions for the distances and different input prices, or fixed distances and different input prices. The only visible change is that, in the first case, the curve for Air Canada is closer to Nordair's rather than Quebecair's. These figures are available upon request to the authors.

result from our example.[15] The presence of economies of scope is important by itself because, as shown by Jara-Díaz (1982b) and Jara-Díaz and Basso (2003) on technological grounds, incentives to merge might appear because of economies of spatial scope even under constant returns to density/disaggregated scale; note that, as emphasised by Basso and Jara-Díaz (2003), this would require $RTD = 1$ and $RTS > 1$, which is analytically impossible. Note also that, although Quebecair has higher returns to density than Air Canada, it presents smaller returns to spatial scope for the same values for *PS*. This further shows that, while returns to density favour the presence of economies of scope, the former do not determine the latter completely.

Conclusions on industry structure regarding Canadian airlines should be taken with care, since we do not have the variance-covariance matrix of the parameter estimates and, therefore, we can not calculate t-statistics for our scope calculations. Nevertheless, the results obtained permit a new interpretation of the process of consolidation of regional airlines in Canada. We have concluded that, irrespective of density and possible increasing *RTD*, the firms with small networks present increasing returns to spatial scope, which makes it cost convenient to expand their networks. While returns to density provide a direct incentive to merge for firms serving the same or highly overlapped network (parallel merging, as *RTD* is calculated explicitly with fixed networks), *SC* provides incentives to merge for firms with non or minimum network overlapping. The relevant question then is: what were the options open to regional carriers? A series of statements in the 1960s, particularly 1969's regional carrier policy, confined the regional carriers' networks to their respective geographical areas with minimum overlap, showing that the potential for parallel merging was limited. This makes our finding of increasing returns to spatial scope a relevant cost explanation for the consolidation observed afterwards, as it would be a means to enlarge the network size in order to take advantage of these economies.

Finally, regarding the issue of how to incorporate changes in average distances, we want to stress that what is important is to make the assumptions explicit when analysing network growth. As in the case of the flows, where we are explicitly imposing that on average they do not change, researchers should be clear on what happens with distances (something dependent not only on the magnitudes of the flows, as discussed). In this

[15] It is not true though that for this cost function *any* airline with a small network will have higher returns to scope than an airline with a large network. It is easy to construct examples where, even with twenty *PS*, an airline has returns to scope smaller than Air Canada's. Thus nothing is defined *a priori*, showing that the results are simply a reflection of what happened in this industry in 1980.

sense, we do not deem it necessary to commit to a specific alternative *a priori*. Whenever possible both the constant and variable average distance cases should be explored but, in some cases, data may preclude it. Which alternative is more "correct" will depend not only on how industry behaved in the past, but also on what one wants to study. Our finding of little difference when using one or the other alternative shows that increasing average distances because of network growth have little impact on the existence of economies of spatial scope in this case.

4.0 Synthesis, Conclusions and Comments

Transport cost functions have to be estimated using aggregate descriptions of product. The question is how these functions can be used to study the potential advantages of expanding the network served. This type of analysis, complementary to the calculation of *RTD* (which has been shown to correspond to scale in a strict sense), has been done in the literature by means of the *RTS* index. But *RTS* has been shown to prevent a meaningful industry structure analysis regarding optimal network size, mainly due to the constant density condition (Basso and Jara-Díaz, 2002). Following Jara-Díaz, Cortés, and Ponce (2001), we have proposed here that we should analyse network expansions by looking at the true product and its translation into a vector of aggregates. Two facts emerged: first, when the network grows new outputs (OD flows) are produced, which means that economies of scope should be calculated. Second, as aggregates are functions of the true product, adding new OD pairs translates into changes of the aggregates, which makes the orthogonal partitions required for scope calculation look simply like evaluations of the cost function in some other values of the aggregates, which have to be calculated. The method presented in this paper rests on a procedure to find the aggregates' counterparts of the partitions in order to use estimated cost functions to calculate economies of scope.

When dealing with empirically estimated cost functions, OD flows are unknown and some condition should be imposed in order to deal with the magnitude of the new flows, that is, those added after the network expansion. We found it reasonable to replace the problematic constant density condition that underlies *RTS* by the constancy of the average OD flow of each type. This was shown to provide enough information to calculate the new aggregates needed to estimate economies of scope. Our proposal is based upon this calculation. The method proposed for economies of spatial scope can not only replace *RTS* in its intended task

— analysing the cost convenience of network expansion — but also encompass an intuitively correct but economically incomplete previous explanation for network growth, namely that by increasing their networks firms feed more traffic into the links and, therefore, partially exploit economies of density. Also, the method can be applied to an estimated translog cost function, not particularly suited to calculate scope but quite useful to calculate other indexes and to impose regularity conditions.

In order to illustrate the method, we presented an application using the results of a published study on Canadian airlines. The calculation of economies of scope emerged very neatly and the case presented fulfilled very well the objective of illustrating the potentials of the method: knowing some information that goes little beyond what is needed for the calculation of *RTS*, the degree of economies of spatial scope can be estimated precisely and unambiguous conclusions regarding network size can be obtained.

The specific conclusions obtained on industry structure regarding Canadian airlines with the new method should be treated with care due to the lack of statistical analysis. Nevertheless, the results obtained permit a new interpretation for the process of consolidation of regional airlines in Canada. We have concluded that, irrespective of density and possible increasing *RTD* (in itself an incentive to merge), the firms with small networks present increasing returns to spatial scope, which makes it cost convenient to expand their networks. Our analysis with the small firms having similar flow levels as the larger firm, and the fact that regional carriers had non overlapping networks, suggest that increasing returns to spatial scope was a major cost saving force driving the merging process. As stated earlier, economies of spatial scope have been suggested previously as an explanation for merging and the emergence of alliances but they have not been calculated from an estimated cost function with aggregate output until now.

A more complete study of industry structure would require a detailed analysis of both economies of density/disaggregated scale and economies of scope, involving different configurations, sizes and number of firms. Just as an example, if a firm that exhibits increasing returns to density/scale ($RTD > 1$) and constant returns to spatial scope ($SC = 0$), decided to exploit its economies of scale/density by increasing output within its fixed network, it might happen that $SC = 0$ turned into diseconomies of scope ($SC < 0$). Note also that economies of *spatial* scope, as calculated here, do not involve changes in proportion between the different services offered (output mix). Since we are situated in a multiproduct setting that includes not only space, industry structure analysis should consider variations of this proportion between different types of services or flows. This can be done through sensitivity analyses, as in the previous section,

but more importantly through the examination of aggregate product-specific economies of scale, cost complementarity between aggregate product types, and economies of scope between different types of services (the type of economies of scope usually calculated in the literature). The fact that space is the most important specific dimension of transport production makes the role of the network particularly important but should not hide the importance of other dimensions.[16] It is also important to recall that this is an analysis in terms of a cost function only, as *RTS* was; potential advantages on the demand side are, therefore, yet to be considered.

The method presented here highlights the importance and usefulness of viewing transport processes in terms of the true product, even if one cannot work empirically with that notion. This is particularly relevant when making economic inferences irrespective of the specific form that the aggregate description of product takes. From this viewpoint, it is necessary to emphasise that what we have presented is a general approach to deal with network expansions, namely the calculation of spatial scope. The specific procedure for the actual calculation of *SC* will depend on the type of aggregates used in the cost function specification, as shown. The analysis of how to proceed with other aggregates and attributes is part of future work.

As a summary, viewing aggregate transport cost functions from the viewpoint of the true transport product has originated both the method presented here to calculate spatial scope and the approach by Jara-Díaz and Cortés (1996) to calculate returns to density/scale correctly. This implies that a fairly complete and meaningful analysis of a transport industry structure can be performed from aggregated transport cost functions provided they are correctly interpreted. Both methods can be applied to most published studies, which motivates a re-evaluation of their policy conclusions. Future studies can take advantage of the econometric advances that have generated many reliable cost functions for different transport industries. We believe that our approach permits a richer analysis.

References

Baltagi, B. H., J. M. Griffin, and D. P. Rich (1995): "Airline Deregulation: The Cost Pieces of the Puzzle," *International Economic Review*, 36, 245–58.

[16]In this sense, we focused on the aspects of Gillen *et al.*'s article that were useful for our application. They indeed analyse matters related to output mix such as cost complementarities between services.

Basso, L. J. and S. R. Jara-Díaz (2002): *Requiem for Returns to Scale with Variable Network Size in Transport Industry Structure Analysis*, Working Paper, Civil Engineering Department, Universidad de Chile, http://www.cec.uchile.cl/~dicidet/lbasso/RequiemEconomiesScale.pdf.

Baumol, W. J., J. C. Panzar, and R. D. Willig (1982): *Contestable Markets and the Theory of Industry Structure*, Harcourt, Brace and Jovanovich, New York.

Braeutigam, R. R. (1999): "Learning about Transport Costs," in Gomez-Ibañez, J., W. B. Tye, and C. Winston (eds.), *Essays in Transportation Economics and Policy*, Brookings Institution Press, Washington DC, 57–97.

Brueckner, J. K. and P. T. Spiller (1994): "Economies of Traffic Density in the Deregulated Airline Industry," *The Journal of Law and Economics*, 37, 379–413.

Caves, D. W., L. R. Christensen, and M. W. Tretheway (1984): "Economies of Density versus Economies of Scale: Why Trunk and Local Airline Costs Differ," *Rand Journal of Economics*, 15, 471–89.

Filippini, M. and R. Maggi (1992): "The Cost Structure of the Swiss Private Railways," *International Journal of Transport Economics*, 19, 307–27.

Gagné, R. (1990): "On the Relevant Elasticity Estimates for Cost Structure Analysis of the Trucking Industry," *The Review of Economics and Statistics*, 72, 160–4.

Gillen, D., T. H. Oum, and M. Threteway (1985): *Airline Cost and Performance: Implications for Public and Industry Policies*, Center for Transportation Studies, University of British Columbia, Canada.

Gillen, D., T. H. Oum, and M. Tretheway (1990): "Airlines Cost Structure and Policy Implications," *Journal of Transports Economics and Policy*, 24, 9–34.

Griliches, Z. (1972): "Cost Allocation in Railroad Regulation," *Bell Journal of Economics*, 3, 26–41.

Hurdle, G. T., R. I. Johnson, A. S. Toskow, G. J. Warden, and M. A. Williams (1989): "Concentration, Potential Entry and Performance in the Airline Industry," *Journal of Industrial Economics*, 38, 119–39.

Jara-Díaz, S. R. (1982a): "The Estimation of Transport Cost Functions: A Methodological Review," *Transportation Review*, 2, 257–78.

Jara-Díaz, S. R. (1982b): "Transportation Product, Transportation Function and Cost Function," *Transportation Science*, 16, 522–39.

Jara-Díaz, S. R. (1988): "Multioutput Analysis of Trucking Operations using Spatially Disaggregated Flows," *Transportation Research B*, 22, 159–71.

Jara-Díaz, S. R., P. Donoso, and J. Araneda (1991): "Best Partial Flow Aggregation in Transportation Cost Functions," *Transportation Research B*, 25, 329–39.

Jara-Diaz, S. R. and C. Cortés (1996): "On the Calculation of Scale Economies from Transport Cost Functions," *Journal of Transport Economics and Policy*, 30, 157–70.

Jara-Díaz, S. R., C. Cortés, and F. Ponce (2001): "Number of Points Served and Economies of Spatial Scope in Transport Cost Functions," *Journal of Transport Economics and Policy*, 35, 327–41.

Jara-Díaz, S. R. and L. J. Basso (2003): "Transport Cost Functions, Network Expansion and Economies of Scope," *Transportation Research E*, 39, 271–88.

Keeler, J. P. and J. P. Formby (1994): "Cost Economies and Consolidation in the U.S. Airline Industry," *International Journal of Transport Economics*, 21, 21–45.

Kirby, M. G. (1986): "Airline Economies of Scale and the Australian Domestic Air Transport Policy," *Journal of Transport Economics and Policy*, 20, 339–52.

Kumbhakar, S. C (1992): "Allocative Distortions, Technical Progress and Input Demand in U.S. Airlines: 1970:1984," *International Economic Review*, 33, 723–37.

Oum, T. H. and M. W. Tretheway (1990): "Airline Hub and Spoke System," *Transportation Research Forum Proceedings*, 30, 380–93.

Oum, T. H. and Y. Zhang (1991): "Utilisation of Quasi-Fixed Inputs and estimation of Cost Functions," *Journal of Transports Economics and Policy*, 25, 121–34.

Oum, T. H. and Y. Zhang (1997): "A Note on Scale Economies in Transport," *Journal of Transport Economics and Policy*, 31, 309–15.

Oum, T. H., J.-H. Park, and A. Zhang (2000): *Globalization and Strategic Alliances: the Case of the Airline Industry*, Pergamon, Elsevier Science.

Panzar, J. and R. Willig (1977): "Economies of Scale in Multioutput Production," *Quarterly Journal of Economics*, 91, 481–93.

Panzar, J. (1989): "Technological Determinants of Firm and Industry Structure," in Schmalensee, R. and R. Willig (eds.), *Handbook of Industrial Organization*, North-Holland, Amsterdam, 4–59.

Small, K. A. (1992): *Urban Transportation Economics*, Harwood Academic Publishers.

Spady, R. and A. F. Friedlaender (1978): "Hedonic Cost Functions For the Regulated Trucking Industry," *Bell Journal of Economics*, 9, 159–79.

Winston, C. (1985): "Conceptual Developments in the Economics of Transportation: An Interpretive Survey," *Journal of Economic Literature*, 23, 57–94.

Xu, K., C. Windle, C. Grimm, and T. Corsi (1994): "Re-evaluating Returns to Scale in Transport," *Journal of Transport Economics and Policy*, 28, 275–86.

Ying, J. S. (1992): "On Calculating Cost Elasticities," *The Logistics and Transportation Review*, 28, 231–5.