Transportation Research Part E 60 (2013) 13–26

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

Explaining changes and trends in the airline industry: Economies of density, multiproduct scale, and spatial scope

Sergio R. Jara-Díaz *, Cristián E. Cortés, Gabriela A. Morales

Civil Engineering Department, Universidad de Chile, Casilla 228-3, Santiago, Chile

ARTICLE INFO

Article history:
Received 17 December 2012
Received in revised form 12 July 2013
Accepted 14 September 2013

Keywords:
Air transport
Network size and shape
Route structure
Scale economies
Spatial scope

ABSTRACT

Changes in the shape and size of airline networks have not been explained clearly from a cost perspective based on the finding of increasing returns to density for given route structures and constant returns to scale for variable network size. We reassessed the estimates of these economies by using new scale and scope indices, finding savings due to changes in route structures and various types of economies of spatial scope not previously calculated: network size, trunk-local services and domestic-international services. Results contribute new insights on the role of cost incentives in the observed changes and trends in the airline industry.

1. Introduction

Most of the empirical discussions on the organization of the airline industry have been conducted with the help of empirical estimations of cost functions, from which economies of density (*RTD*) and economies of scale (*RTS*) can be obtained. *RTD* addresses output growth with a constant network size and route structure, whereas *RTS* considers changes in the network size. Using these concepts, potential cost savings arising from output and/or network growth have been studied in the specialized literature, where most authors have found increasing *RTD* and nearly constant *RTS*. According to the definitions of both indices, previous studies have concluded that there are cost advantages arising from increasing the flow while holding the network size constant (density of traffic is increased) and that there are no cost advantages for firms in operating larger networks. However, the behavior observed in the industry has not followed these trends. After deregulation (in the US first and then in the rest of the world), the concentration and network size of the industry have increased through mergers, acquisitions, and alliances. Mergers are often viewed as a way to facilitate marketing management by the airlines to build anti-competitive monopoly power in pricing. Although the search for higher profits commanded by demand has likely had a role in this evolution, cost studies have attempted to explain network growth along with the emergence of hub-and-spoke operating structures through *RTD* and *RTS*.

On the pure cost side, some authors have argued that the observed behavior of airlines can be understood "as an attempt to exploit economies of traffic density" (Brueckner and Spiller, 1994) and that this behavior occurs "in spite of constant returns to scale/network size" because "the addition of a station to a hub and spoke system can result in economies" (Oum and Tretheway, 1990), which might occur because the density over the existing spokes could increase. As stated by Basso and Jara-Díaz (2005), "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 that serve the same set of nodes but, as found in every econometric study

* Corresponding author. Tel.: +56 2 29784380; fax: +56 2 26894206.
E-mail addresses: jaradiaz@ing.uchile.cl (S.R. Jara-Díaz), ccortes@ing.uchile.cl (C.E. Cortés), gamorale@ing.uchile.cl (G.A. Morales).

1366-5545/$ - see front matter
http://dx.doi.org/10.1016/j.tre.2013.09.003

that has considered a network variable, expanding the network is costly and this is not considered in the density justification for network growth''. As shown by Kumbhakar (1990) and Liu and Lynk (1999), economies of density were important before deregulation but are less important in the post-deregulation period. This observation has been reinforced by Swan (2002), who concluded that the main source of economies of density, namely, aircraft size, has not played as relevant a role as the change in the route structure in recent decades. This conclusion coincides with the findings of Wei and Hansen (2003) regarding the importance of frequency over aircraft size to accommodate larger flows. Other authors have argued that an alternative explanation for merging or the formation of alliances has been economies of spatial scope (see, for example, Hurdle et al., 1989; Oum et al., 2000). The levels of efficiency that can be achieved in procurement and sharing facilities with larger airlines have often been viewed as important. However, costs may rise as unions have more concentrated power with a single carrier, and wage rates may shift toward the higher level of the two merging companies.

Thus, the changes observed in the airline industry have been interpreted and explained using estimated cost functions in different ways, generating interesting discussions and debate. In this paper, we first review the empirical literature in air transport from 1984 to 2012, synthesizing the analysis and conclusions regarding industry structure based on estimations of *RTD* and *RTS*. We then present three new indices that have been developed in the literature to replace the prevailing scale measures: *RTD′*, the corrected version of *RTD*, which examines costs as flows expand holding the route structure constant; the multioutput degree of scale economies (*S*), which considers costs as flows expand allowing for the readjustment of the route structure; and economies of spatial scope (*SC*), which examines the advantages or disadvantages of serving two distinct sets of flows. We find that $RTD < RTD' < S$ and that there are various types of economies of spatial scope, including network size, trunk-local services, and domestic-international services. These results advance beyond previous analyses of the role of cost incentives behind the observed changes and trends in the airline industry. Route forms and network shapes and sizes could also be driven by demand characteristics, although these effects are not analyzed here. Strategic interaction among competitors also contributes to the structure of air networks (Oum et al., 1995).

In the next section, the concepts of *RTD* and *RTS* are presented, and their application to the analysis of the air industry is summarized. The new indices developed in the transport economics literature are then introduced and justified. The use of these new indices to re-examine each case is presented in Section 3. Finally, a synthesis and conclusions are offered in Section 4.

2. Cost functions in air transport

2.1. Empirical estimates of returns to density and returns to scale: a synthesis

The multioutput degree of economies of scale, *S*, is defined by the maximal proportional expansion of output *Y* that is feasible after a proportional expansion of inputs *X*, such that a value of *S* larger than, equal to, or smaller than one implies increasing, constant, or decreasing returns to scale, respectively (Panzar and Willig, 1977). Although *S* is defined based on the technology, *S* can be calculated from the cost function $C(w, Y)$, which represents the minimum expenditure necessary to produce *Y* at input prices *w*, as

$$S = \frac{C(Y)}{\sum_i y_i \frac{\partial C}{\partial y_i}} = \frac{1}{\sum_i \eta_i}, \tag{1}$$

where η_i is the elasticity of *C* with respect to the *i*th output. Input prices are omitted for notational simplicity. Note that $S > 1$ implies that a proportional expansion of *Y* induces an increase in cost by a smaller proportion.

The large size that the output vector *Y* achieves in the transport sector (i.e., flows in very many origin–destination (OD) pairs) precludes its direct use in the empirical work, as noted by various authors. Thus, cost functions must be estimated using aggregate output descriptions, $\widetilde{Y} = \{\tilde{y}_h\}$, such as ton-kilometers or seat-kilometers, and so-called *attributes*, such as average distance or load factor. When a network size variable *N* is used, empirical studies of transport industries distinguish between two concepts of '*scale*': returns to density (*RTD*) and returns to scale (*RTS*), as first proposed by Caves et al. (1984). In the former case, the network is assumed to be fixed as output increases, i.e. traffic density increases. In the latter case, both output and network size increase, with traffic density remaining unchanged. *RTD* is calculated as the inverse of the sum of a subset of the cost-output elasticities (this subset varies from study to study and has thus become a source of ambiguity). In *RTS*, the elasticity of the network size is also included in the calculation. Note that *RTD* represents an attempt at capturing what *S* actually measures, i.e. the change in cost as output expands. *RTS* was designed to address the cost effects of network expansion, i.e., a variation in *N*.

In the airline market literature, most authors estimate a cost function considering similar outputs and attributes in the specification. In Table 1, we summarize the literature covering 30 years during and after deregulation of US, Canadian, and European markets. Output and the so-called attributes are assumed to be exogenous in these studies. Overall, a translog form has been used to specify the cost function and, in all cases, the number of points served, *PS*, is used as the indicator of network size *N*. The typical output aggregates are of the form flow × distance, such as revenue passenger mile (*RPM*) and revenue ton miles (*RTM*), as discussed in Oum and Zhang (1991), Liu and Lynk (1999), and Creel and Farell (2001). In some cases, the aggregates are grouped into one variable called multilateral output index (Caves et al., 1984; Kumbhakar, 1990; Windle, 1991; Oum and Yu, 1998). The hedonic output (grouping each aggregate with its associated attributes) introduced