



**HAL**  
open science

# Who is My Main Competitor? Measuring Airline Pairwise Competitive Intensity Based on Route Supply Strategies

Qian Liu, Paul Rochet, Chantal Roucolle

► **To cite this version:**

Qian Liu, Paul Rochet, Chantal Roucolle. Who is My Main Competitor? Measuring Airline Pairwise Competitive Intensity Based on Route Supply Strategies. 2024. hal-04728168

**HAL Id: hal-04728168**

**<https://hal.science/hal-04728168v1>**

Preprint submitted on 9 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Who is My Main Competitor? Measuring Airline Pairwise Competitive Intensity Based on Route Supply Strategies

Qian LIU<sup>a</sup>, Paul ROCHET<sup>a</sup>, Chantal ROUCOLLE<sup>a</sup>

<sup>a</sup>*Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, France*

---

## Abstract

Classical approaches to the study of airline competition involve either assessing market concentration or studying rivalry dynamics, through adaptive responses to competition. However, these methods often lack precision at the individual level or reach the limits of computational complexity. To leverage the advantages of both methods, we discuss various measures of competitive intensity between two airlines using bivariate statistical analysis, based on their flight supply strategies across European routes. First, four pairwise symmetric indices are presented and compared. The weighted Jaccard index is identified as the best measure and used to examine the European domestic market in the first and third quarters of 2023. The results show the consistency of the index with real competitive situations. In addition, an asymmetric version of the weighted Jaccard index is introduced for further research and discussion.

*Keywords:* Pairwise competition, Asymmetric index, European airlines, Competitive intensity, Airline supply strategy

---

## 1. Introduction

The intensity of competition in the air transport market has been increasing with the deregulation of markets, beginning with the US Deregulation Act in 1978, extended by the European Single Sky in 1992, or by the ASEAN (Association of Southeast Asian Nations) Single Sky in 2023. In this highly competitive environment, airlines must adapt their strategies to maintain their profitability and performance levels. Strategies include collaborations, mergers, business model convergence, pricing adjustments, and scheduling changes, all aimed at reducing competition or regaining market share from direct competitors. Sun et al. (2024) propose a comprehensive survey on airline competition research, categorizing airline competition studies into ten aspects: route entry and exit decisions, service quality, network competition, concentration and alliances, among others.

Since 2005, the US domestic market has been characterized by several mergers among main airlines, resulting in five competitors sharing over 80% of traffic. This consolidation aims to reduce market competitive intensity. Empirical literature indicates that competitive intensity is one of the most important determinants of airline decisions, although its impact remains unclear and may depend on the characteristics of the markets studied, and the variables of interest. For instance, considering scheduled time as a measure of quality, Brueckner et al. (2021), Prince and Simon (2015) and Mazzeo (2003) find that increased competition may lead airlines to implement larger flight buffers, as competition heightens their focus on maintaining on-time performance. Similarly, Greenfield (2014) finds that competition improves on-time performances, Fan (2019) shows that more competition leads to shorter scheduled flight block times, and Roucolle et al. (2020b) indicate a nonlinear relationship between competition and excess travel time. They find that extreme market conditions, such as monopolies or highly competitive markets, are characterized by longer excess travel times, corresponding to higher inefficiency. In addition, the nature of competitors may also affect airline decisions and performances. Bubalo and Gaggero (2015) show that the presence of low-cost carriers at European airports leads to reduced delays. Moreover, the presence of Southwest Airlines on the US domestic market influences the behaviors of its competitors, as shown in Goolsbee and Syverson (2008), or Ciliberto and Tamer (2009).

---

*Email address:* qian.liu@enac.fr (Qian LIU)

From a broader perspective, competitor identification is a multifaceted topic. For instance, Bergen and Peteraf (2002) address it through a broad-based managerial approach, while Yu et al. (2016) suggest that the perceived environmental uncertainty can influence firm's perception of its competitors. Although the identification of competitors remains complex, airline competition fits better with Cournot competition (Brander and Zhang, 1990) with homogeneous products and competing on flight capacity. When choosing flight capacity, airlines should first decide which route to serve, then at what frequency and with which aircraft. In this context, the airlines with the largest network overlap level should be regarded as the closest and most direct competitors.

Airlines consider their competitors in their decision-making process. In this context, we argue that a pairwise competition index should be relevant to capture the adaptive behaviors of airlines. We investigate airline pairwise competition and address two questions: How can competition between two airlines be measured? Is pairwise competition symmetric, given the heterogeneity of airlines' importance in the market? We propose statistical indicators, based on route frequencies, to provide a quantitative measure of the intensity of competition between two airlines. Our study offers three significant contributions to the literature on airline competition.

Firstly, we propose an index that provides a quantitative measure of airline pairwise competition. Previous studies assess competition either at an aggregate market level, often referred to as "market concentration", or from dynamic perspective. The former approach lacks precision regarding individual behavior. Specifically, market concentration indices, such as the Herfindahl-Hirschman Index (HHI) (Bromberg, 2023), and the Concentration Ratio (CR), can sometimes lead to different outcomes (Lijesen et al., 2002). On the other hand, studies focusing on dynamic processes use behavioral approaches to derive equilibria from airlines' best response functions (Lijesen and Behrens, 2017) or use demand functions (Flores-Fillol, 2009), (Ciliberto et al., 2021), focusing on dynamic responses to rivalry rather than providing immediate quantitative assessments of the degree of competition. In addition, this approach has limited applications due to its computational complexity. Our study integrates the strengths of both approaches: the precision of pairwise analysis and the ability to summarize competition information between airlines through an index. We contribute to the existing literature on airline competition by proposing an index, based on the adaptation of several statistical indices, to measure the intensity of competition between two airlines.

Secondly, we propose straightforward statistical indicators based on airlines' network structures and market shares, which are highly interpretable. These indicators use information about the airlines' flight frequencies on the routes they operate. They measure the outcome of competitive strategies, rather than the competitive decisions themselves. While airline competition is often described in the literature as multimarket contact (Bernheim and Whinston, 1990) (Bilotkach, 2011), our approach focuses specifically on frequency competition. We argue that the proposed indices are consistent with the characteristics of airline competition, seen as Cournot competition. Consequently, our approach does not require categorizing airline types to compare them, allowing the assessment of the competitive intensity between full-service and low-cost carriers based on their route supply strategies.

Thirdly, we include asymmetric indicators to reflect the reality of the air transport market and provide deeper insights into the competitive dynamics faced by airlines. Indeed, Chen and Hambrick (1995) posit that the intensity of competition exhibits an asymmetry, differing between smaller and larger companies. This concept was further developed and substantiated by Chen (1996), who measured and illustrated competitive asymmetry based on market commonality and resource similarity.

Starting from the symmetric case by assuming equal competitive pressure between two airlines, we propose four candidate indices, using normalized frequencies, based on the concept of bivariate analysis. We then compare their different characteristics. In addition, competition between airlines includes both market share and network structure. Therefore, we compare the distributions of normalized airline frequencies on the routes they operate, and ultimately select the weighted Jaccard index as the optimal measure of competition intensity between a pair of airlines. The selected Jaccard index is then calculated for individual airlines and at the group level. The group level approach is in line with the previous literature, which suggests that there is no competition between coordinated or integrated airlines (see Forbes and Lederman (2009), Gillen et al. (2015), and Roucolle et al. (2020a)). Finally, an asymmetric index is presented for discussion, which is intended to better capture the asymmetric nature of competition between airlines or airline groups.

We find that the weighted Jaccard index effectively measures pairwise competitive intensity, by capturing the essential features of airline competition in a simple and interpretable way. It aligns with the real-world measurement of network overlap between airlines. Furthermore, the asymmetric version of the weighted Jaccard index offers a more detailed and realistic measure of competition, that facilitates the identification of an airline's main competitors.

The remainder of this paper is structured as follows. In Section 2, we construct symmetric indices focusing on supply strategy, and propose four symmetric candidate indices to find a robust and relevant pairwise competition measure. In Section 3, we present the results both at the individual airline level and at the group level, with the weighted Jaccard index selected as the most relevant one. In Section 4, the definition for the asymmetric index is presented, as an extension of the weighted Jaccard index. Numerical results for airline groups are then given to illustrate how the index could be interpreted. Section 5 discusses the limitations of the proposed indices. Finally, Section 6 concludes the paper.

## 2. Symmetric indices

In our research, we aim to answer the following question: How can we develop a quantitative index to measure the intensity of competition between two airlines? This index should be able to provide a quick insight into the competitive situation between airlines.

It is crucial to distinguish between our focus on “competitive intensity” and the usual notion of “level of competitiveness” (Bhawsar and Chattopadhyay, 2015). The latter usually refers to the overall extent of competition present in a market, considering factors such as the number and strength of competitors, market share distribution, pricing strategies, product differentiation, and barriers to entry. In contrast, “competitive intensity” focuses on the degree of aggressiveness or fierceness of competition between specific competitors. It pays special attention to the intensity of rivalry and the active pursuit of market share, customers.

This section outlines various approaches to conceptualizing a symmetric index, which measures the intensity of competition between two airlines. The competitive intensity is considered symmetric if the airline size (total flight frequency) is ignored, and the focus is instead on the flight allocation strategy. The indices are calculated using the flight frequencies. Indeed intra-European flights predominantly use narrowbody aircraft, so seat capacity is nearly proportional to flight frequency. The use of flight frequency as a measure of supply capacity aligns with capturing airlines’ behavior, as flight frequency captures key competitive elements such as route dominance, capacity allocation, and market share. The observed frequencies represent the airlines’ best responses in equilibrium (Lijesen and Behrens, 2017). Later in Section 4, we propose an alternative viewpoint by introducing an asymmetric index that accounts for heterogeneity in airline sizes.

### 2.1. Concept of a “good” index

A robust and relevant index should satisfy the following properties:

- **No competition:** If a pair of airlines have no common routes, they are not competing against each other. Ideally, the value of the competitive intensity index should be zero or negligible.
- **The strongest competition:** If the networks of two airlines completely overlap and the airlines employ identical flight allocation strategies (i.e. the same proportion of flights relative to their total frequency on each route), they are considered to be in the most intense competition. In this scenario, the competitive intensity index should reach its maximum value.
- **Prioritization of route overlap:** Route overlap increases the likelihood of head-to-head competition between airlines for passengers traveling to the same destination. Therefore, we believe that competition among airlines primarily revolves around their flight networks rather than frequencies on specific routes, with airlines with greater route overlap facing more intense competition, regardless of their respective market share.

### 2.2. Index 1: correlation coefficient

Since studying the behavior between two airlines fits the concept of studying a bivariate relationship, which is a specific case in multivariate analysis (Timm, 2002), a natural approach is to examine the correlation of route supply between two airlines of interest (or called focal airlines). By analyzing supply information based on the spatial composition of each airline pair (i.e. flight frequency on different routes), rather than focusing on each airline’s share in a given market as HHI does, this approach measures pairwise competition and takes into account the differences in

the spatial structure of each airline pair’s network. This allows for a targeted examination of behavioral differences between the two airlines.

The first suggested index establishes the correlation of airline supply in a certain geographic region  $R$ , such as an ensemble of countries, cities, or airports. Given that airlines compete primarily on routes, we define  $R = \{r_1, r_2, \dots, r_n\}$  as the geographic area in which they operate. The cardinality  $|R| = n$ , means that there are  $n$  routes in  $R$ , with each route represented by  $r_i$ . The frequency on a route refers to the total number of round-trip flights, including outbound and return flights.  $f^a(r_i) \in \mathbb{Z}^+$  is the number of flights supplied by airline  $a$  on route  $r_i \in R$ , and

$$\bar{f}^a = \frac{1}{|R|} \sum_{r_i \in R} f^a(r_i)$$

denotes the average flight frequency of airline  $a$ . The covariance between two airlines  $a$  and  $b$  within the region  $R$  is defined in (1). The competition index is the associated Pearson correlation coefficient, shown in (2).

$$\sigma^{(a,b)} = \frac{1}{|R|} \sum_{i=1}^n (f^a(r_i) - \bar{f}^a)(f^b(r_i) - \bar{f}^b) \quad (1)$$

$$I_1^{(a,b)} = \frac{\sigma^{(a,b)}}{\sqrt{\sigma^{(a,a)}} \sqrt{\sigma^{(b,b)}}}. \quad (2)$$

The index  $I_1^{(a,b)}$  lies within  $[-1, 1]$ . Theoretically, the sign of a linear correlation should convey meaningful information: a positive correlation might indicate competition between airlines, with a value of 1 denoting the most intense competition; a negative correlation suggests the absence of competition; and a correlation close to zero implies independent supply strategies. However, the calculation of the correlation coefficient includes zeros. When there are no shared routes between two airlines, the resulting index value can be negative rather than zero. Additionally, negative values may arise even in cases of substantial route overlap between two airline networks, which poses a challenge to the interpretation of these values. Another drawback is that the Pearson correlation is highly sensitive to outliers. If outliers occur on two airlines’ disjoint routes, the values may be distorted and not accurately represent the overall relationship. Alternative coefficients such as Spearman’s or Kendall’s rank correlation, which do not rely on a linear relationship, lead to very similar results.

### 2.3. Index 2: cosine similarity

Cosine similarity is calculated by dividing the dot product of the vectors by the product of their magnitudes, as shown in (3). It measures the cosine of the angle between two vectors. For example, two proportional vectors have a cosine similarity of 1, orthogonal vectors have a similarity of 0, and opposite vectors have a similarity of -1. It is also widely used for measuring the similarity between two documents (Singhal et al., 2001).

$$I_2^{(a,b)} = \frac{f^a \cdot f^b}{\|f^a\| \|f^b\|} = \frac{\sum_{i=1}^n f^a(r_i) f^b(r_i)}{\sqrt{\sum_{i=1}^n f^a(r_i)^2} \cdot \sqrt{\sum_{i=1}^n f^b(r_i)^2}} \quad (3)$$

The index  $I_2^{(a,b)}$  lies in  $[0, 1]$ , as all vectors  $f^a = (f^a(r_1), \dots, f^a(r_n))^t$  and  $f^b = (f^b(r_1), \dots, f^b(r_n))^t$  hold positive values. The index equals zero when the airline pair has no common routes, with the two frequency vectors being orthogonal. The value of 1 indicates the most intense competition. Additionally, cosine similarity measures the angle between two vectors, making it insensitive to airline sizes by construction. Note that centered cosine similarity, where the attribute vectors are normalized by subtracting the vector means (e.g.,  $f^a - \bar{f}^a$ ), is equivalent in value to the Pearson correlation coefficient.

### 2.4. Index 3: based on the dissimilarity of airline supply distributions

The dissimilarity between populations can be measured using an Euclidean distance metric or other definitions based on entropy and discrimination (Rao, 1982). The competition between two airlines can also be interpreted as negatively correlated with the distance between them. In other words, the greater the distance between the airlines (indicating that they are “farther apart” from each other), the lower the competitive intensity.

Consider the probability mass functions for airlines  $a$  and  $b$

$$p^a(r_i) = \frac{f^a(r_i)}{f^a(R)} \quad , \quad p^b(r_i) = \frac{f^b(r_i)}{f^b(R)} \quad (4)$$

where  $\sum_{i=1}^n p^a(r_i) = 1$ ,  $\sum_{i=1}^n p^b(r_i) = 1$  with  $n = |R|$ .  $p^a$  and  $p^b$  represent the flight frequencies normalized by the total frequency of airlines  $a$  and  $b$ . They correspond to the probabilities assigned by distributions  $p^a$  and  $p^b$  to the route  $r_i$ . The probability measures  $p^a$  and  $p^b$  can be viewed as the strategies of airlines, representing their flight allocation on each route. The pairwise competitive intensity can therefore be measured by a decreasing function of their distribution distance  $d(p^a, p^b)$ .

Possible metrics between distributions include the Chi-Squared Distance, Hellinger Distance (Hellinger, 1909), Earth Mover's Distance (also known as Wasserstein Distance (Vaserstein, 1969)), Kolmogorov-Smirnov Distance (An, 1933), although the simplest for our purposes is arguably the Total Variation Distance. It is worth noting that the entropy-based Kullback-Leibler Divergence (Kullback and Leibler, 1951) is essentially used for distributions with a common support (in our case, the routes operated by an airline), making it unsuitable for our situation.

The Total Variation Distance (TVD) between distributions  $p^a$  and  $p^b$  can be calculated as defined in (5), where  $p^a(r_i)$  and  $p^b(r_i)$  are defined in (4). The summation is taken over all possible outcomes in  $R$ .

$$\text{TVD}(p^a, p^b) = \frac{1}{2} \sum_{i=1}^n |p^a(r_i) - p^b(r_i)| \quad (5)$$

Since  $\text{TVD}(p^a, p^b) \in [0, 1]$ , to gauge the index within  $[0, 1]$  for better interpretability, we define the index  $I_3^{(a,b)}$  in (6).

$$I_3^{(a,b)} = 1 - \text{TVD}(p^a, p^b) \quad (6)$$

The index  $I_3^{(a,b)}$  lies within  $[0, 1]$ , with a value of zero indicating no common routes between airlines, as their associated probabilities are disjoint, and a value of 1 signifying the most intense competition. It quantifies the similarity between airlines based on their flight allocation strategy, specifically the proportion of allocated flights on shared routes.

#### 2.5. Index 4: weighted Jaccard similarity

The Jaccard index (Jaccard, 1912), also referred to as the Jaccard similarity coefficient, is a statistic used to assess the similarity between finite sample sets. For  $X$  and  $Y$  two finite sample sets, the index is defined as the size of the set intersection divided by the size of the union in (7).

$$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} = \frac{|X \cap Y|}{|X| + |Y| - |X \cap Y|} \quad (7)$$

It is worth noting that in the literature on air transport networks, the Jaccard index can be used to measure changes in airport networks over two time periods (Bringmann et al., 2018) (Li et al., 2023), or to assess airport brand perception by measuring social network cluster similarity of airports (Gitto and Mancuso, 2019).

Therefore, an index of competitive intensity for airlines can naturally be constructed using the concept of the Jaccard index, as

$$I_{\text{Jaccard}}^{(a,b)} = \frac{N(a \cap b)}{N(a \cup b)} \quad (8)$$

with  $N(a \cap b)$  denoting the number of common routes between the two airlines, and  $N(a \cup b)$  representing the total number of routes covered by at least one of the two airlines. However, this index  $I_{\text{Jaccard}}^{(a,b)}$  exclusively quantifies the network overlap between airlines  $a$  and  $b$ , without taking into account the frequency-based competition on shared routes.<sup>1</sup> Therefore, we consider the weighted Jaccard similarity between two airline strategies  $p^a, p^b$  defined in (9) as a competitive intensity index.

<sup>1</sup>For example, if two airlines both operate the same routes (so they have the same overlap), then their competition will change due to fluctuations in flight frequency, causing the index to adjust accordingly.

$$I_4^{(a,b)} = \frac{\sum_{i=1}^n \min(p^a(r_i), p^b(r_i))}{\sum_{i=1}^n \max(p^a(r_i), p^b(r_i))} \quad (9)$$

The index  $I_4^{(a,b)}$  lies in  $[0, 1]$ , and equals zero when there are no common routes between the airlines. It measures the weighted overlap rate by frequency, normalized by each airline's total frequency capacity. A potential drawback compared to the previous indices is that it exhibits minimal variation, which may make the results less distinguishable and hinder outlier detection.

Although they are based on somewhat different ideas, the two indices  $I_3^{(a,b)}$  and  $I_4^{(a,b)}$  are in fact linked by the relation

$$I_4^{(a,b)} = \frac{I_3^{(a,b)}}{2 - I_3^{(a,b)}}. \quad (10)$$

This becomes obvious when noticing

$$\sum_{i=1}^n \min(p^a(r_i), p^b(r_i)) + \sum_{i=1}^n \max(p^a(r_i), p^b(r_i)) = 2$$

and

$$\text{TVD}(p^a, p^b) = 1 - \sum_{i=1}^n \min(p^a(r_i), p^b(r_i)).$$

This explains their similar behavior and near perfect collinearity, as shown in Section 3.2, e.g. in Tables 5 and 6.

## 2.6. Concluding remarks

To define a symmetric pairwise competition index between airlines based on network flight frequencies, we focus on the following four aspects:

- (i) The index should be easily interpretable within its value range.
- (ii) The absence of overlapping routes between the two airlines should unambiguously signify an absence of competition.
- (iii) Since it is symmetric by construction, the index should not be affected by the total frequency of airlines, but only by the flight distributions within their networks.
- (iv) The presence/absence of two airlines on a given route is regarded as a main factor in measuring the intensity of competition. Therefore, the index should prioritize airlines' common networks rather than the relative frequencies on the shared routes, although ideally both should be considered.

Following this paradigm, the Pearson correlation coefficient  $I_1$  appears to be the most promising in terms of interpretability. It has a broader range of values with three discernible states, each with a specific meaning: negative (collaboration), near zero (indifference) and positive (competition). In practice however, even for airlines with no common routes, it remains mostly close to zero, thereby failing the second requirement. On this account, the cosine similarity  $I_2$  seems better suited, since a minimal theoretical value of zero (due to flight frequency vectors being positive) is equivalent to the absence of common routes between the two airlines. Both the Pearson correlation coefficient and cosine similarity are conceived to be invariant to scale changes, thereby fulfilling the third requirement. However, they fail to address consideration (iv), which emphasizes the focus on shared networks, as they all appear to be strongly affected by changes in shared route frequency trends.

Indices  $I_3$  and  $I_4$  are extremely similar and nearly collinear. In the situation of low competition, the values of both indices are close to zero. While they fulfill all four conditions, a preliminary normalization of the airline frequencies

$f^a$  and  $f^b$ , justifying the use of distributions  $p^a$  and  $p^b$ , is required to achieve scale invariance (iii). A summary of the properties of the four symmetric indices is given in Tab. 1.

Table 1: Properties of the four symmetric indices.

	$I_1$	$I_2$	$I_3$	$I_4$
<b>(i)</b>	✓	✓	✓	✓
<b>(ii)</b>	×	✓	✓	✓
<b>(iii)</b>	✓	✓	✓	✓
<b>(iv)</b>	×	×	✓	✓

### 3. Numerical results for symmetric indices

In this section, we present the data used to analyze and illustrate the performance of each index proposed in Section 2. We use airline route frequencies aggregated for Q1 and Q3 of 2023. By comparing the four candidate indices, we identify the weighted Jaccard index (index 4) as the most pertinent symmetric index. Using index 4, we first calculate values for individual airlines identified by their airline codes<sup>2</sup>, regardless of ownership relations, to provide a snapshot of the index’s performance. Given that airlines are not always financially or strategically independent and cannot always be regarded as competitors, we then perform a calculation at airline group level to account for common ownership.

#### 3.1. Data

We analyze data sourced from the Official Airline Guide (OAG), a database providing details on global scheduled flights. This dataset includes various route and flight characteristics, such as operating carriers, published carriers, departure and destination airports, route distance in kilometers, departure and arrival times, flight duration, available seats per flight and flight date.

We extract one-way, non-stop direct passenger flights in 2023, specifically during the summer months (the third quarter or Q3: July, August, September) and the winter months (the first quarter or Q1: January, February, March). We aggregate round-trip flight frequencies for each season (three months) to analyze round-trip flights on airport origin-destination (OD) for the domestic European market. We choose to study the flights within the 41 Eurocontrol member states (see Table 2) because they operate under the Single European Sky project, which ensures the same freedom to set operations.<sup>3</sup>

The choice between studying city pairs or airport pairs as OD is always debated. Brueckner et al. (2014) has suggested a methodology of grouping airports into city pairs, but concluded that the substitutability of airports depends on trip characteristics, which are not always discernible from available data. Given that there is no firm conclusion in the literature, we use airport pairs as OD units.

We selected 20 focal airlines (Table 3) based on their flight frequencies in the summer (Q3) of 2023. After selection, the same airlines are also analyzed for Q1. First, we filtered out routes with strictly less than 10 flights, indicating that these routes are operated fewer than 10 times (equivalent to less than 3 flights per month) by any airline in Q3. This accounts for 4.1% of routes in Q3 and 15.4% of routes in Q1. We believe this disparity between Q1 and Q3 can be attributed to airlines reducing leisure travel supply during the winter, which is in particular the case for low-cost carriers and charters. After excluding routes with fewer than 10 occurrences in a quarter, there are 351 operating carriers in Q3, with 10,159 routes and a total route frequency of 2,400,233. In Q1, 334 airlines operate 7,254 routes, with a total of 1,640,609 flights. Each route is considered undirected, meaning that the OD pair A-B is equivalent to B-A.

Subsequently, we identified the top 15 operating carriers with the highest total flight frequency in Q3. We also included five important subsidiaries associated with these top 15 airlines to evaluate the proposed indices. This practice is common in the literature and assumes no competition between coordinated or integrated airlines (see Forbes and

<sup>2</sup>Individual airlines are identified by their unique IATA or ICAO codes in OAG.

<sup>3</sup><https://www.europarl.europa.eu/factsheets/en/sheet/133/air-transport-single-european-sky>



Lederman (2009), Gillen et al. (2015), and Roucolle et al. (2020a)). These 20 selected airlines account for 50.10% of the total frequency in Q3 and 53.00% of that in Q1.

Finally, we grouped the 20 selected airlines based on the criterion that airlines with the same holding belong to the same group, whether due to mergers, acquisitions, or the formation of subsidiaries. This makes a total of 10 groups (see Table 4). By gathering airlines in this way, we can analyze competition between airline groups and avoid calculating index values for non-competing entities. We then calculate the four indices at airline level and compare the Jaccard index at both airline and group levels.

Table 2: Definition of Europe: 41 Eurocontrol Member States.

Albania	Germany	Poland
Armenia	Greece	Portugal
Austria	Hungary	Republic of Moldova
Belgium	Ireland	Romania
Bosnia & Herzegovina	Italy	Serbia
Bulgaria	Latvia	Slovakia
Croatia	Lithuania	Slovenia
Cyprus	Luxembourg	Spain
Czech Republic	Malta	Sweden
Denmark	Monaco	Switzerland
Estonia	Montenegro	Turkey
Finland	Netherlands	Ukraine
France	North Macedonia	United Kingdom
Georgia	Norway	

Source: <https://www.eurocontrol.int/our-member-and-comprehensive-agreement-states>

### 3.2. Index validation

We conclude that the four proposed indices measure competitive intensity from different perspectives and focuses. Despite this, they demonstrate consistency with one another, supported by the identification of the same top 5% largest values. Ultimately, we select the weighted Jaccard index  $I_4$  as the most relevant option.

*Correlation among indices* We initiated our analysis by calculating pairwise indices for the Q3 data of the 20 selected airlines in Table 3, resulting in 190 values per index. Next, we computed the correlation coefficients between each pair of indices, both with and without the extreme values. The indices maintain the same notations as previously stated:  $I_1$  stands for the Pearson correlation coefficient,  $I_2$  for the cosine similarity,  $I_3$  for the total variation distance between distributions, and  $I_4$  for the weighted Jaccard similarity. Table 5 presents the linear correlation coefficients for each pair of indices, and Table 6 shows the linear correlation coefficients excluding the top 5% largest values, which are considered extreme. These extreme values represent airline pairs engaged in the most intense competition. Comparing the two tables, we note that extreme values have a significant influence on the correlation coefficients. We observe that the index  $I_1$  differs significantly from the indices  $I_2$  and  $I_3$ , as evidenced by the low correlation when extreme values are excluded. The index  $I_2$  shows a high correlation with the indices  $I_3$  and  $I_4$ . The strong correlation between the indices  $I_3$  and  $I_4$  is due to the exact relation  $I_4 = I_3/(2 - I_3)$  previously highlighted in (10). The similarity is even emphasized in Table 6 where excluding the largest values makes them almost linear, in view of the Taylor expansion around the origin:  $I_4 \approx I_3/2$  as  $I_3 \approx 0$ . Moreover, the results demonstrate the consistency of the indices with each other, as the four indices share common extreme values<sup>4</sup>

<sup>4</sup>(‘FR’, ‘VY’), (‘FR’, ‘LS’), (‘MAY’, ‘W4’), (‘VY’, ‘TO’), (‘PC’, ‘TK1’), (‘W6’, ‘RR’), (‘WA’, ‘KL’). Additionally, indices  $I_1$  and  $I_2$  have two other identical extreme values: (‘U2’, ‘AF’), (‘U2’, ‘LS’), which differ from those of  $I_3$  and  $I_4$ : (‘U2’, ‘FR’), (‘FR’, ‘MAY’).

Table 3: The 20 Selected Airlines.

Airline code	Airline name	Frequency (in %)
U2	EasyJet	6.84%
FR	Ryanair	6.59%
TK	Turkish Airlines	4.46%
MAY	Malta Air	3.90%
LH	Deutsche Lufthansa AG	3.63%
VY	Vueling Airlines	2.62%
BA	British Airways	2.55%
AF	Air France	2.36%
PC	Pegasus Airlines	2.22%
W4	Wizz Air Malta	1.58%
LS	Jet2.com	1.58%
W6	Wizz Air	1.57%
SK	SAS Scandinavian Airlines	1.48%
OS	Austrian Airlines AG dba Austrian	1.40%
TK1	AnadoluJet	1.39%
WA	KLM Cityhopper	1.33%
RR	Buzz	1.27%
KL	KLM Royal Dutch Airlines	1.19%
TO	Transavia.com France	1.09%
IB	Iberia	1.05%
Q3 Total frequency:		Total:
2400233		50.10%

Source: OAG Analyser.

Note: They are ranked in descending order of flight frequency. The first 15 represent the highest frequency carriers operating during the Q3 (July, August, and September) of 2023. The remaining 5 are subsidiaries included to evaluate the efficiency of the proposed indices.

Table 4: Airline Groups of the 20 Selected Airlines.

Airline Group	Group codes	Airline IATA codes	Frequency (in %)
Ryanair Group	FR	FR, MAY, RR	11.76%
EasyJet Group	U2	U2	6.84%
International Air-line Group (IAG)	IAG	VY, BA, IB	6.22%
Air France - KLM	AFKLM	AF, TO, WA, KL	5.96%
Turkish Airlines	TK	TK, TK1	5.85%
Lufthansa Airlines	LH	LH, OS	5.03%
Wizz Air Group	W6	W4, W6	3.15%
Pegasus Airlines	PC	PC	2.22%
Jet2.com	LS	LS	1.58%
SAS Group	SAS	SK	1.48%
Q3 Total frequency:	2400233	Total:	50.10%

Source: OAG Analyser.

Table 5: Correlation Coefficient Matrix of Indices Based on Q3 Data of 20 Airlines.

	$I_1$	$I_2$	$I_3$	$I_4$
$I_1$	1.000			
$I_2$	0.972	1.000		
$I_3$	0.932	0.973	1.000	
$I_4$	0.944	0.982	0.995	1.000

Note: There are 190 points in total. The indices used are as follows.  $I_1$ : Pearson correlation coefficient;  $I_2$ : cosine similarity;  $I_3$ : distance between distributions;  $I_4$ : weighted Jaccard similarity.

Table 6: Correlation Coefficient Matrix of Indices Based on Q3 Data of 20 Airlines, excluding the largest 5% extreme values.

	$I_1$	$I_2$	$I_3$	$I_4$
$I_1$	1.000			
$I_2$	0.690	1.000		
$I_3$	0.595	0.918	1.000	
$I_4$	0.593	0.917	<b>0.999</b>	1.000

Note: There are 190 points in total. The indices used are as follows.  $I_1$ : Pearson correlation coefficient;  $I_2$ : cosine similarity;  $I_3$ : distance between distributions;  $I_4$ : weighted Jaccard similarity.

*Diverse focuses of the indices* However, it is worth noting that these indices have distinct focuses in measuring competitive intensity, as evidenced by the different extreme values identified. For instance, the pair ('U2', 'AF') shares 38 routes, while ('FR', 'MAY') shares 378. The pair ('U2', 'AF') is identified as extreme by  $I_1$  and  $I_2$  because they have a strong positive correlation in their frequency allocation trends on these routes, despite having fewer route overlaps compared to other pairs. In contrast,  $I_3$  and  $I_4$  prioritize route overlap over correlation of frequencies. This explains why ('FR', 'MAY') is detected by  $I_3$  and  $I_4$  but not by  $I_1$  and  $I_2$ .

The distinct focuses of these indices are also illustrated in Figure 1, where we calculated the values of the four indices for each pair of airlines within the same group (Table 4), resulting in 15 pairs. We identified two patterns of airline pairs sharing overlapping routes:

1. **Same sign for all indices:** For pairs such as ('WA', 'KL'), ('AF', 'KL'), ('BA', 'IB'), and ('VY', 'IB'),  $I_1$  has the same positive sign as the other three indices. This indicates that these subsidiaries within the same group are imperfect substitutes, with flights allocated interchangeably among them.
2. **Opposite or negligible  $I_1$  values:** For pairs like ('FR', 'MAY'), ('FR', 'RR'), ('MAY', 'RR'), and ('W4', 'W6'), the  $I_1$  values are negligible or negative, whereas the other three indices indicate a contrasting trend. This suggests that, despite some overlap in their route network structures, the frequencies on these routes are allocated in opposing directions: one rise and the other one decrease. This phenomenon implies that these subsidiaries are strategically reducing their flights on the same routes to avoid price wars or overcapacity, thereby maximizing group profitability and market share.

From the correlation table excluding extreme values (Table 6), we observe that  $I_2$ ,  $I_3$ , and  $I_4$  measure similar aspects with a quasi-linear relationship, while  $I_1$  provides a different perspective. Although the combined analysis using  $I_1$  and one of the other three indices can offer a nuanced view of the relationship, the  $I_1$  index can be biased by extreme frequency values of routes not in common. Moreover, our emphasis lies more on route overlap than on frequency allocation trends, as explained in Section 2.1. Therefore, we prefer to use indices  $I_3$  or  $I_4$ , which are more suitable for measuring competitive intensity.

*Distributions of the indices* We explore the distributions of the four symmetric indicators with a boxplot in Figure 2. The presence of numerous outliers above the upper whisker suggests a heavy-tailed distribution of the index values, reflecting the oligopoly situation in aviation. The small size of the box indicates that most data points are tightly clustered around 0. This clustering is characteristic of airlines either avoiding competition in many cases or experiencing sparse data, often due to operating on a limited number of routes within a vast geographical area.

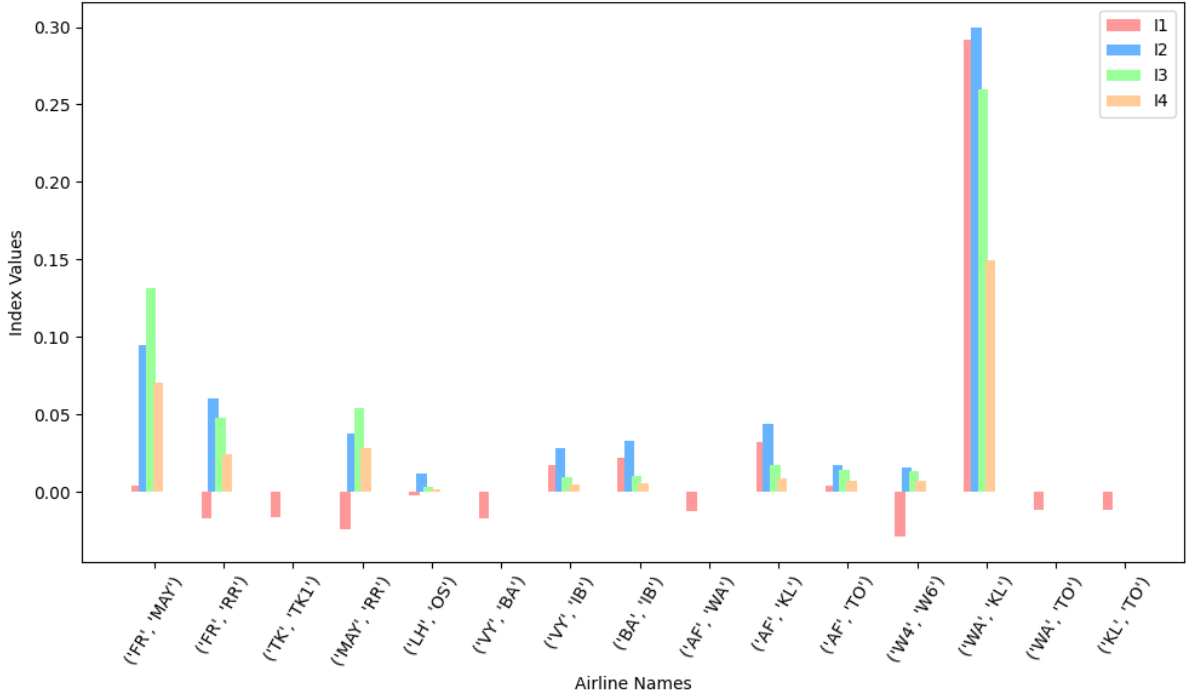


Figure 1: Index values for airline pairs, with each pair belonging to the same group. 2023 Q3 data.

*Concluding remarks* Based on the comparative analysis of the four indices considered,  $I_1$  appears as the weakest one while the other three tend to behave very similarly. In the sequel, we choose to focus the analysis on  $I_4$  only. This decision is based on the strong explainability of the weighted Jaccard similarity for competitive relationships between airlines, as well as for its natural asymmetric extension which will be discussed in Section 4. Besides, note that the absolute size of the values themselves does not have significant interpretative meaning. Instead, these values are used for comparison purposes. For instance, if the value for one pair of airlines is larger than the value for another pair, it indicates that the competition between the pair with the larger value is more intense.

### 3.3. Results for airlines

First, we calculate indices for individual airlines (Table 3). Each operating carrier, distinguished by a unique airline code in OAG, is here treated as an independent airline, regardless of its group affiliation or ownership relations. For example, AF and KL are considered separate entities in this analysis, despite their operational integration under the Air France-KLM group. Prior to calculations, we observed that airlines adjust their networks according to seasons. This adjustment is evidenced by the change in the number of routes between summer and winter, with airlines operating 10,159 routes in summer and reducing to 7,254 routes in winter.

Additionally, the analysis of the index values indicates that seasonality shapes the market competition structure, leading to shifts in the main competitor pairs. This can be observed from the changes in Figure 3. Using the weighted Jaccard index ( $I_4$ ) as the measure of pairwise competitive intensity, the figure is a scatter plot with summer index values on the x-axis and winter index values on the y-axis. The points on the blue diagonal line ( $y = x$ ) represent pairs with the same competitive intensity in summer and winter. Points above the diagonal indicate that these pairs experience more intense competition in winter than in summer, while points below the line indicate greater competition in summer than in winter. For example, we can conclude that the pair ('WA', 'KL') exhibits a significant route overlap during summer. However, in winter, they do not share the same network. This observation also reveals a distinctive “collaboration” pattern, which we will discuss further in Section 3.4.

Furthermore, from the perspective of an individual airline, its primary direct competitor varies over time. This variation is illustrated by the fluctuating competitive index between Ryanair and other airlines (Figure 4). As previ-

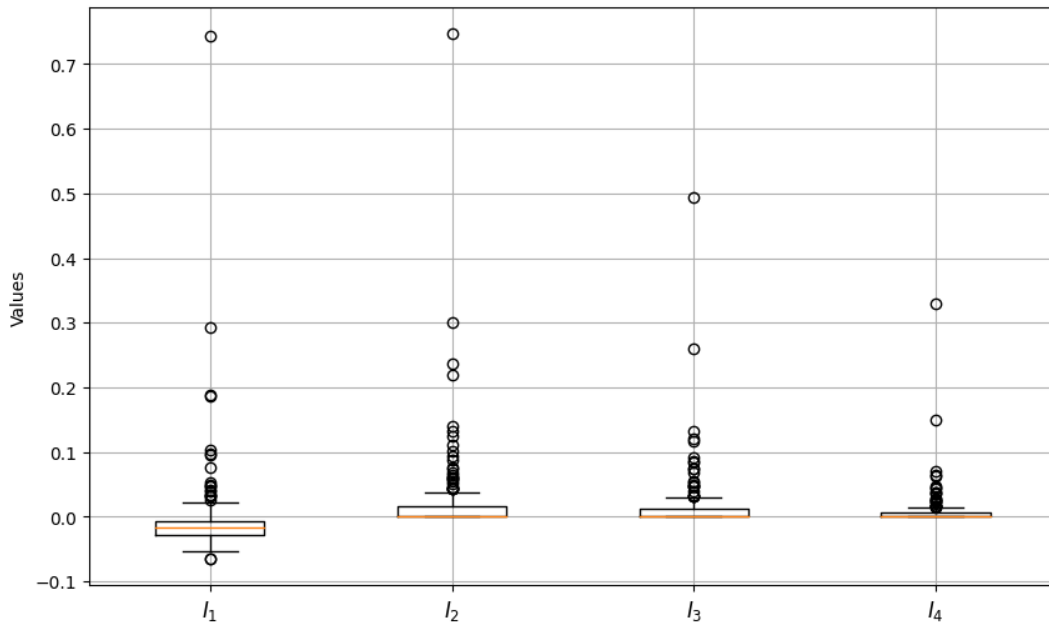


Figure 2: Box plot of the four proposed symmetric indices, based on Q3 data of 20 airlines.

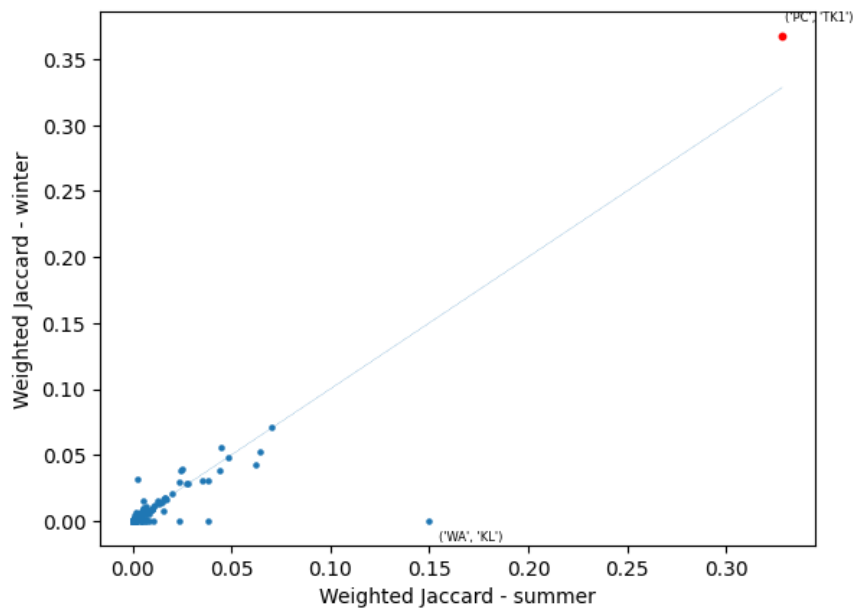


Figure 3: Competitive intensity between summer and winter for all pairs of individual airlines as operating carriers.

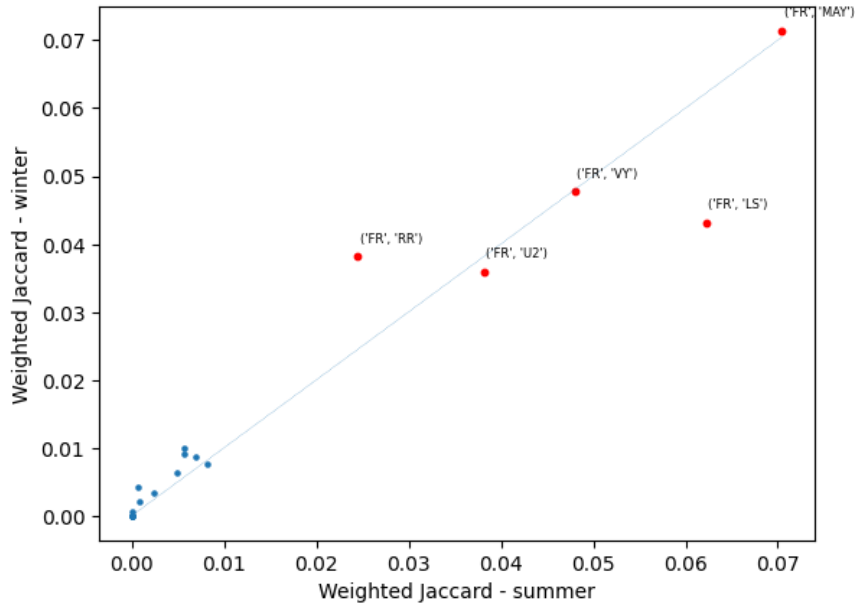


Figure 4: Competitive intensity between summer and winter for pairs of competitors solely involving Ryanair.

ously explained, the scatter plot points do not perfectly align with the diagonal line ( $y = x$ ), indicating that competitive relationships change with the seasons. For example, we observe that Malta Air, a subsidiary of Ryanair, exhibits the highest route overlap with Ryanair, with the pair ('FR', 'MAY') showing the largest values during both winter and summer seasons. This aligns with Ryanair's strategy of transferring routes to and from Malta to this new airline, with plans for further network expansion (Hudson, 2019). During summer 2023, Jet2 emerges as Ryanair's second main competitor, evidenced by the pair ('FR', 'LS') having the second largest index value. However, in winter, Vueling becomes Ryanair's second main competitor. Indeed, we observe that Jet2 reduced its number of routes from 375 in Q3 to 185 in Q1, while overlapping routes decreased from 137 to 92. In contrast, Vueling maintained a similar number of routes and overlapping routes with Ryanair during the two quarters (278 and 251 routes in Q3 and Q1; 83 and 85 common routes in Q3 and Q1). These findings underscore the dynamic nature of competitive relationships among airlines.

### 3.4. Results for airline groups

We then compute index values for airline groups, after aggregating airlines with the same holding company (Table 4). Calculating at the group level allows to exclude airline pairs with the same ownership, ensuring that we focus on genuine competitors. Indeed, calculating indices between group members can be misleading in practice, as they generally do not view each other as competitors. This group approach enhances the reliability and realism of the results.

Consistent with our previous observations for airlines, we notice seasonal variations in the competitive dynamics among airline groups. Illustrated in Figure 5, the group pair Easyjet and Air France-KLM ('U2', 'AFKLM'), exhibits reduced competition in winter, yet experiences substantial route overlap in summer. In contrast, the group pair ('TK', 'PC') consistently engages in intense competition, regardless of the time of year. This persistence may be due to the fact that the group Turkish Airlines ('TK') strategically established AnadoluJet ('TK1') as a competitive countermeasure against the threat posed by Pegasus Airlines ('PC'), as shown by Tasci and Yalcinkaya (2015). This explains their extensive route overlap leading to the high level of competition.

From the perspective of an airline group, its primary direct competitors can be identified using this competitive intensity index. Similar as for airlines, the competitive intensity also varies seasonally at group level. For instance, Figure 6 illustrates this concept for the Ryanair group. We can conclude that Ryanair does not compete with the groups TK and PC, as the two indices equal to zero. Its competitive relationships remain stable with the groups AFKLM,

SAS, and LH, while exhibiting slight seasonal variations with W6 and U2. Notably, Ryanair competes more intensively with the group LS in the summer of 2023 than in the winter.

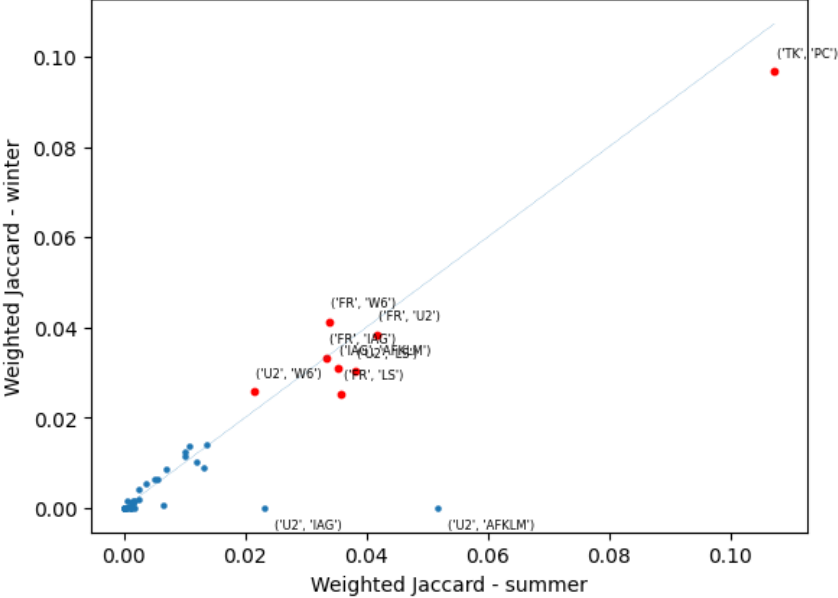


Figure 5: Competitive intensity between summer and winter for all pairs of airline groups.

Another noteworthy observation is the distinctive “collaboration” pattern among airlines. Generally, airlines within the same group “collaborate” by integrating distinct networks to achieve economies scope, as exemplified by the merger of Air France and KLM in 2004. However, we note in Figure 4 that the pairs (‘FR’, ‘MAY’) and (‘FR’, ‘RR’) have significant route overlap, even though they are both within the Ryanair group. This occurs because the holding company FR maintains the routes of its subsidiaries. A similar conclusion can be drawn for the pair (‘WA’, ‘KL’) in Figure 3, and for the pair (‘MAY’, ‘RR’) in Figure 1.

Apart from revealing the seasonal changes in competitive relationships among airline groups (Figure 5), the relationships of airlines within a specific time period can be graphically represented by drawing a network of airline relations. An airline competition network is illustrated in Figure 7, using the symmetric pairwise index values as edge weights, where edges with darker colors indicate more intense competition. Therefore, for a focal airline, we can instantly recognize its main direct competitor by identifying the airline with which the index has the largest value. For instance, the primary direct competitor for AF-KLM group is U2. For Q3 2023 data, we can identify two clusters in Figure 7. The first cluster relates to large airline groups. For example, the AF-KLM group notably competes with IAG and U2, as the pairs (‘AFKLM’, ‘IAG’) and (‘AFKLM’, ‘U2’) have larger index values. Similarly, the airline group pairs (‘U2’, ‘LS’), (‘U2’, ‘W6’), (‘U2’, ‘IAG’), (‘FR’, ‘LS’), (‘FR’, ‘U2’), (‘FR’, ‘W6’), and (‘FR’, ‘IAG’) have intense competitive relationships. As for the second cluster, we identified (‘TK’, ‘PC’) as a local or regional competition in Turkey, as they have weak competitive intensity with others but extremely strong competition between them.

**4. Asymmetric index**

In the previous sections, our focus is on comparing symmetric pairwise indices, meaning that the intensity of competition is considered equally strong for company *a* as it is for *b*. This symmetric consideration is suitable when comparing the distribution of flights for each airline. For example, if we believe that competition is better measured by comparing the proportion of flights allocated on a route relative to the airline’s overall capacity, rather than by comparing the absolute number of flights, we can employ this method. In this way, the competitive intensity index

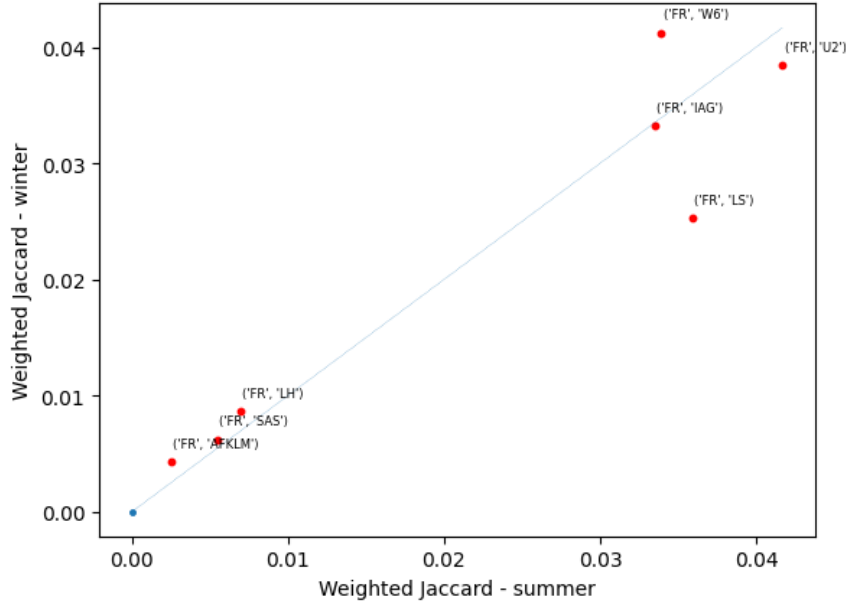


Figure 6: Competitive intensity between summer and winter for pairs of competitors at group level, solely involving Ryanair Group.

could be the same for two large airlines as for one large and one small airline, if they employ the same flight allocation strategies on overlapping routes.

However, it is crucial to acknowledge that competition between two airlines can also be seen as asymmetric, especially when considering airline size. The same situation may be more critical for a small airline than for a large one. This argument is evidenced by Chen and Hambrick (1995). In reality, using a symmetric index, a large airline would always have a small index value when compared with a small airline due to their low route overlap. However, this value does not capture the differences in competition faced by airlines  $a$  and  $b$  given their distinct competitive positions. An asymmetric index can address this. For instance, a smaller airline  $a$  with only 20 routes does not face the same level of competition as a competitor  $b$  operating on 200 routes and sharing common routes. Therefore, the competitive intensity caused by  $a$  to  $b$  is typically low, as a small airline poses little threat to a large airline. Conversely, the competitive intensity of  $b$  applied to  $a$  is much larger, given its stronger market power.

#### 4.1. Asymmetric variation of index 4

To account for the potentially unbalanced intensity of competition between two airlines, the unweighted Jaccard similarity based index can naturally be adapted to an asymmetric version as follows:

$$I_{\text{Jaccard}}^{(a|b)} = \frac{N(a \cap b)}{N(b)}. \quad (11)$$

This index measures the proportion of routes operated by airline  $a$  where airline  $b$  is present, reminiscent of the conditional probability of finding airline  $a$  on routes operated by airline  $b$ . Consistent with the competition paradigm, high values of the index indicate that airline  $a$  tends to focus on routes on which airline  $b$  operates. Since this indicator only accounts for the overlapping routes between the two airlines and not on the frequencies on these routes, we propose the weighted version of the asymmetric Jaccard index, defined by

$$I_4^{(a|b)} := \frac{\sum_{i=1}^n \min(f^a(r_i), f^b(r_i))}{\sum_{i=1}^n f^b(r_i)}. \quad (12)$$



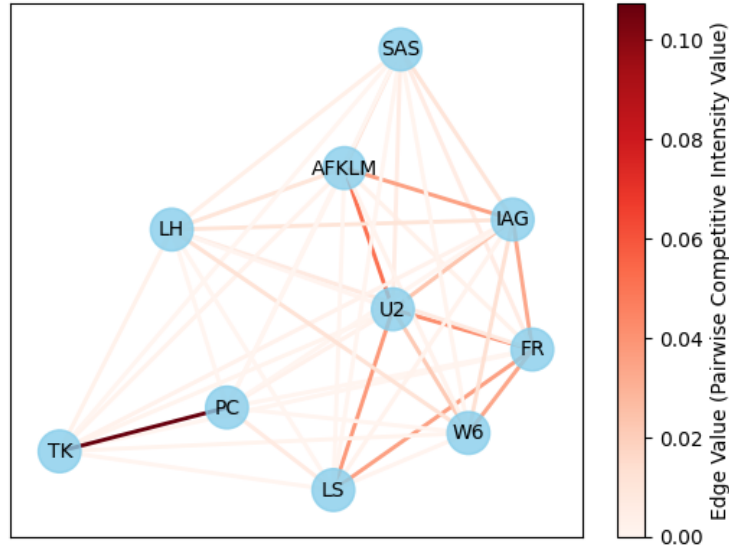


Figure 7: Graphical representation of airline competitive relationships based on the symmetric pairwise competitive intensity index, which is calculated as the weighted Jaccard similarity. Q3 2023 data.

#### 4.2. Numerical results for asymmetric index

Using the asymmetric weighted Jaccard index as defined in (12), we calculate the competitive intensity of all nine competitors of the Ryanair group for the summer and winter of 2023. Results are illustrated in Figure 8 and Figure 9. Specifically, we interpret the results in terms of competition disparity and seasonality.

*Competition disparity* For the summer of 2023, results are shown in Figure 8 with a Dumbbell plot, illustrating the differences between airline groups in various competitive positions. We note that the pair (‘FR’, ‘LS’) exhibits the most significant difference in flight overlapping. For example, the asymmetric index  $I_4^{(LS | FR)}$  measures 0.051, meaning that only 5.1% of FR’s flights are overlapped by LS. Conversely, the index  $I_4^{(FR | LS)}$  measures 0.377, indicating that 37.7% of LS’s flight propositions are also offered by FR. Therefore, the competitive pressure from FR to LS is much more significant than from LS to FR. In other words, in the competition between FR and LS, LS is in a weaker position compared to FR. Moreover, the plot is ranked from the largest difference to the smallest. Therefore, pairs with small differences such as (‘FR’, ‘LH’) and (‘FR’, ‘AFKLM’) tend to have similar flight network sizes. Notably, we also observe that FR has no direct competition with the airline group TK and the group PC, given their null index values. Similar conclusions can be drawn from Figure 9.

*Seasonality* The sum of index values for  $x | FR$  in summer 2023 totals 0.252, indicating that these nine group competitors of Ryanair in total have a supply overlap of 25.2% with Ryanair group. This total is 23.5% in winter. If we plot the index values of each group pair, Figure 10 indicates a slight shift in the competitive relationships between winter and summer. Notably, the third largest direct competitor of Ryanair becomes LS in summer, whereas it was W6 in winter. Moreover, seasonality is depicted in Figure 8 and Figure 9, as the differences in asymmetric index values between airline groups vary across different seasons.

## 5. Discussion

Measuring the competitive intensity between two airlines is a complex process. The following aspects of the previously proposed indices could be better addressed.

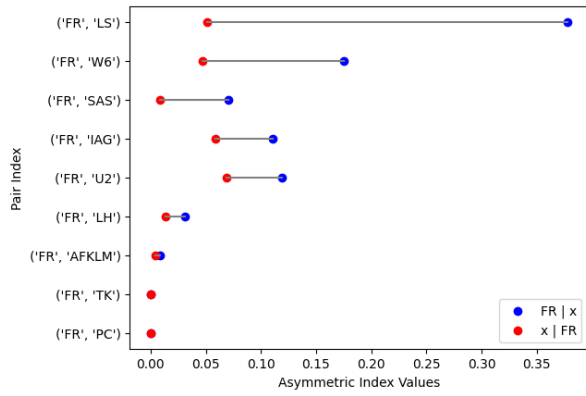


Figure 8: Dumbbell plot of asymmetric competitive intensity index values for airline groups competing with Ryanair in Q3 of 2023. The line between two points shows the difference between their values.

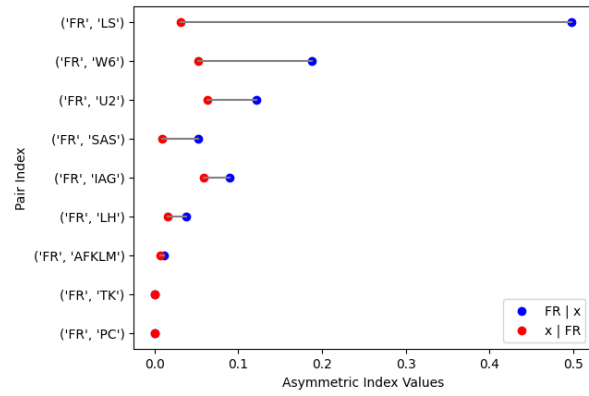


Figure 9: Dumbbell plot of asymmetric competitive intensity index values for airline groups competing with Ryanair in Q1 of 2023. The line between two points shows the difference between their values.

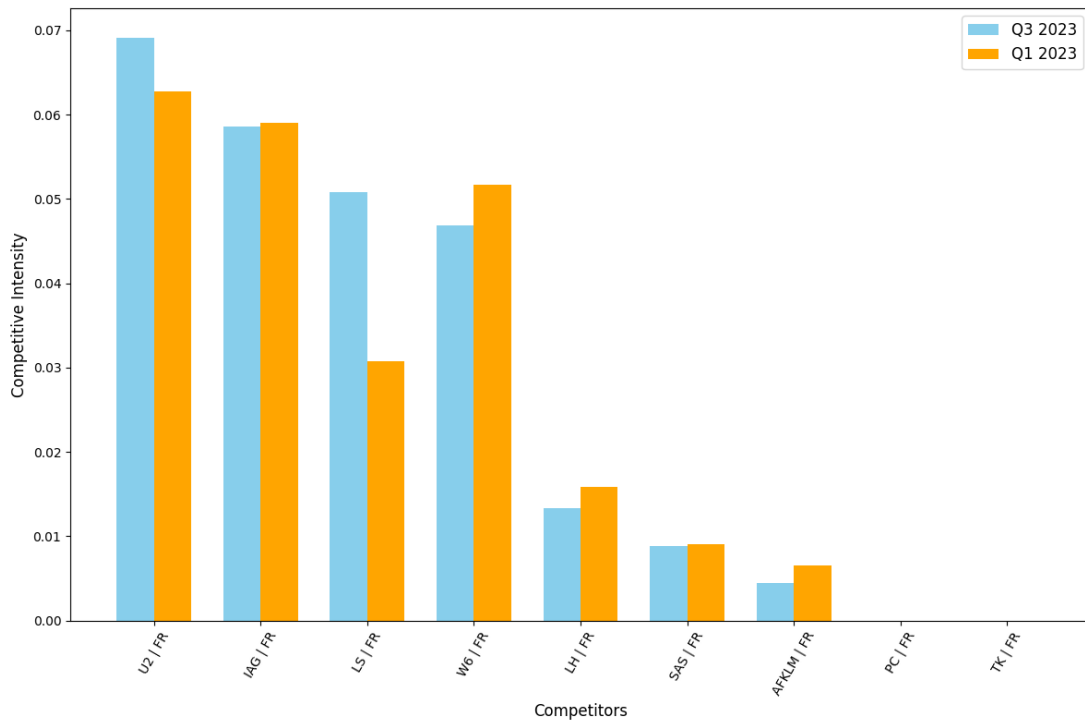


Figure 10: Asymmetric competitive intensity faced by Ryanair group from its competitors in Q1 and Q3 of 2023. Plot is arranged in decreasing order of Q3 values.

### 5.1. Sparsity

Upon reviewing the initial results of the proposed symmetric index  $I_4^{(a,b)}$ , we observed that the index values are sensitive to the sparsity of the airline network. For instance, in Q3 2023, Ryanair ('FR') operated on 1602 routes, while Malta Air ('MAY') operated on 956 routes, resulting in 378 overlapping routes and a weighted Jaccard index of 0.070. In contrast, AnadoluJet ('TK1') had only 165 routes and Pegasus Airlines ('PC') 250 routes, with 100 common routes between them, yielding a weighted Jaccard index of 0.329, which is larger than that of the pair ('FR', 'MAY'). This observation suggests that larger airline networks may perceive lower competition against focal competitors, as competition appears "diluted" due to a lower proportion of overlapping routes relative to the total network size.

Another approach to extend the results for symmetric and asymmetric indices is to aggregate all index values that measure pairwise competitive intensity for a focal airline. This aggregation would provide a measure of the total competition faced by that airline or airline group. However, it's important to scale this measure to ensure comparability across different airlines.

### 5.2. Other drivers of competition

From the perspective of supply, the symmetric index  $I_4^{(a,b)}$  effectively measures competitive intensity between two airlines. However, comprehensive measurement of competition may require a systemic approach that integrates both qualitative and quantitative analyses. The current quantitative approach overlooks aspects of product differentiation, such as variations in aircraft type or onboard services. Therefore, it is recommended to consider both resource similarity and market commonality for a comprehensive analysis, as stated in Chen (1996). The existing indices focus solely on flight frequency data and could benefit from incorporating additional variables to provide a more holistic understanding of competitive dynamics.

Furthermore, the study only examined direct passenger flights, omitting how airlines compete through connecting flights via hubs.

### 5.3. Airline competitive clusters

It is noteworthy that symmetric indices (Figure 7) can be used for the identification of airline competitive clusters. These clusters help to clarify the competitive relationships among airline groups, providing insight into their interactions and strategic positioning. Furthermore, clustering methods such as DBSCAN (Density-Based Spatial Clustering of Applications with Noise) could be used to identify outliers, e.g. clusters containing only one element. These outliers could typically indicate that the airlines share similar flight scheduling strategies.

## 6. Conclusion and perspectives

From a new perspective using bivariate analysis to study competition, we measure airline pairwise competitive intensity with relevant statistical indices. These indices are easy to calculate and provide immediate insights into airline competition in a given market. The final proposed weighted Jaccard index and its asymmetric form can serve as reliable signals for conducting competitor analysis and promptly identifying main competitors for an airline or airline group. By illustrating the results of these competitive indices and their implications, we show how to identify the main direct competitors of an airline or airline group based on market share and network structure. Consequently, we enhance understanding of competition between two airlines as a pair. Besides, we uncover the influence of seasonality on variations in competitive intensity.

We empirically compared the four proposed symmetric indices using airline capacity (flight frequency) data for direct flights within Europe. We show that the four proposed symmetric indices adopt a similar approach to measure pairwise competitive intensity by ignoring the total size of the airline. Even though each has its own strengths and weaknesses, we find that these indices exhibit consistency in terms of extreme values and numerical representation. Ultimately, we choose to retain the weighted Jaccard index as the most relevant symmetric index and propose its asymmetric version, considering the airline total capacity. The use of an asymmetric index allows to analyse and study competition between large and small airlines, which are not directly comparable. Finally, results show that both the symmetric index using weighted Jaccard similarity and its asymmetric counterpart are relevant for instantly measuring the competitive intensity between two airlines.

Our analysis presents significant potential for further research. The proposed indices are valuable for identifying local competition and understanding the competitive relationships among airlines. Future research aims to develop a comprehensive asymmetric index of pairwise competitive intensity by integrating the various factors mentioned in the discussion part, and employing time-series analysis to monitor airline competition over time. This will enhance our understanding of the evolution of airline relationships and competitive dynamics from a statistical perspective.

By incorporating an asymmetric index and analyzing the temporal evolution of competitive intensity, we can explore the air transport market stability and resilience in response to market shocks through spatio-temporal modeling. Authorities can utilize these indices for improved market analysis, competitive benchmarking, and regulatory insights. Airlines can identify direct competitors, allocate resources more effectively, track dynamic competition, and formulate targeted strategies, facilitating informed decision-making.

## Funding

This research was possible thanks to the financial support of the Région Occitanie/Pyrénées-Méditerranée, and the École nationale de l'aviation civile (ENAC).

## References

- An, K., 1933. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'istituto Italiano degli Atti* 4, 89–91.
- Bergen, M., Peteraf, M.A., 2002. Competitor identification and competitor analysis: a broad-based managerial approach. *Managerial and decision economics* 23, 157–169.
- Bernheim, B.D., Whinston, M.D., 1990. Multimarket contact and collusive behavior. *The RAND Journal of Economics*, 1–26.
- Bhawsar, P., Chattopadhyay, U., 2015. Competitiveness: Review, reflections and directions. *Global Business Review* 16, 665–679.
- Bilotkach, V., 2011. Multimarket contact and intensity of competition: evidence from an airline merger. *Review of industrial organization* 38, 95–115.
- Brander, J.A., Zhang, A., 1990. Market conduct in the airline industry: an empirical investigation. *The RAND Journal of Economics*, 567–583.
- Bringmann, K., De Langhe, K., Kupfer, F., Sys, C., Van de Voorde, E., Vanelslander, T., 2018. Cooperation between airports: A focus on the financial intertwinement of European airport operators. *Journal of Air Transport Management* 69, 59–71.
- Bromberg, M., 2023. Herfindahl-Hirschman index (HHI) definition, formula, and example. Investopedia.
- Bruceckner, J.K., Czerny, A.I., Gaggero, A.A., 2021. Airline mitigation of propagated delays via schedule buffers: Theory and empirics. *Transportation Research Part E: Logistics and Transportation Review* 150, 102333.
- Bruceckner, J.K., Lee, D., Singer, E., 2014. City-pairs versus airport-pairs: A market-definition methodology for the airline industry. *Review of Industrial Organization* 44, 1–25.
- Bubalo, B., Gaggero, A.A., 2015. Low-cost carrier competition and airline service quality in Europe. *Transport Policy* 43, 23–31.
- Chen, M.J., 1996. Competitor analysis and interfirm rivalry: Toward a theoretical integration. *Academy of Management Review* 21, 100–134.
- Chen, M.J., Hambrick, D.C., 1995. Speed, stealth, and selective attack: How small firms differ from large firms in competitive behavior. *Academy of Management Journal* 38, 453–482.
- Ciliberto, F., Murry, C., Tamer, E., 2021. Market structure and competition in airline markets. *Journal of Political Economy* 129, 2995–3038.
- Ciliberto, F., Tamer, E., 2009. Market structure and multiple equilibria in airline markets. *Econometrica* 77, 1791–1828.
- Fan, T.P.C., 2019. Schedule creep—in search of an uncongested baseline block time by examining scheduled flight block times worldwide 1986–2016. *Transportation Research Part A: Policy and Practice* 121, 192–217.
- Flores-Fillol, R., 2009. Airline competition and network structure. *Transportation Research Part B: Methodological* 43, 966–983.
- Forbes, S.J., Lederman, M., 2009. Adaptation and vertical integration in the airline industry. *American Economic Review* 99, 1831–1849.
- Gillen, D., Hashemina, H., Jiang, C., 2015. Strategic considerations behind the network–regional airline tie-ups—a theoretical and empirical study. *Transportation Research Part B: Methodological* 72, 93–111.
- Gitto, S., Mancuso, P., 2019. Brand perceptions of airports using social networks. *Journal of Air Transport Management* 75, 153–163.
- Goolsbee, A., Syverson, C., 2008. How do incumbents respond to the threat of entry? Evidence from the major airlines. *The Quarterly Journal of Economics* 123, 1611–1633.
- Greenfield, D., 2014. Competition and service quality: New evidence from the airline industry. *Economics of Transportation* 3, 80–89.
- Hellinger, E., 1909. Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen. *Journal für die reine und angewandte Mathematik* 1909, 210–271.
- Hudson, D., 2019. Malta air officially launched, as Ryanair promises \$1 billion investment. *MaltaToday.com.mt* URL: [https://www.maltatoday.com.mt/news/national/95602/watch\\_malta\\_air\\_officially\\_launched\\_as\\_ryanair\\_promises\\_1\\_billion\\_investment](https://www.maltatoday.com.mt/news/national/95602/watch_malta_air_officially_launched_as_ryanair_promises_1_billion_investment). retrieved 2019-06-11.
- Jaccard, P., 1912. The distribution of the flora in the alpine zone. 1. *New Phytologist* 11, 37–50.
- Kullback, S., Leibler, R.A., 1951. On information and sufficiency. *The Annals of Mathematical Statistics* 22, 79–86.
- Li, S., Xu, G., Zhou, Y., 2023. How air transport networks respond to long-lasting disruptive events like COVID-19: The first step toward long-term resilience. *Transportation Research Part A: Policy and Practice* 177, 103836. URL: <https://www.sciencedirect.com/science/article/pii/S0965856423002562>, doi:<https://doi.org/10.1016/j.tra.2023.103836>.

- Lijesen, M., Behrens, C., 2017. The spatial scope of airline competition. *Transportation Research Part E: Logistics and Transportation Review* 99, 1–13.
- Lijesen, M.G., Nijkamp, P., Rietveld, P., 2002. Measuring competition in civil aviation. *Journal of Air Transport Management* 8, 189–197.
- Mazzeo, M.J., 2003. Competition and service quality in the us airline industry. *Review of Industrial Organization* 22, 275–296.
- Prince, J.T., Simon, D.H., 2015. Do incumbents improve service quality in response to entry? evidence from airlines' on-time performance. *Management Science* 61, 372–390.
- Rao, C.R., 1982. Diversity and dissimilarity coefficients: a unified approach. *Theoretical population biology* 21, 24–43.
- Roucolle, C., Seregina, T., Urdanoz, M., 2020a. Measuring the development of airline networks: Comprehensive indicators. *Transportation Research Part A: Policy and Practice* 133, 303–324.
- Roucolle, C., Seregina, T., Urdanoz, M., 2020b. Network development and excess travel time. *Transport Policy* 94, 139–152.
- Singhal, A., et al., 2001. Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.* 24, 35–43.
- Sun, X., Zheng, C., Wandelt, S., Zhang, A., 2024. Airline competition: A comprehensive review of recent research. *Journal of the Air Transport Research Society*, 100013.
- Tasci, D., Yalcinkaya, A., 2015. A new business model in the airline industry: Airline within airline model and a comparison through the case of anadolujet. *ESKISEHIR OSMANGAZI UNIVERSITESI IIBF DERGISI-ESKISEHIR OSMANGAZI UNIVERSITY JOURNAL OF ECONOMICS AND ADMINISTRATIVE SCIENCES* 10.
- Timm, N.H., 2002. *Applied multivariate analysis*. Springer.
- Vaserstein, L.N., 1969. Markov processes over denumerable products of spaces, describing large systems of automata. *Problemy Peredachi Informatsii* 5, 64–72.
- Yu, C.L., Wang, F., Brouters, K.D., 2016. Competitor identification, perceived environmental uncertainty, and firm performance. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration* 33, 21–35.