Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
An empirical assessment of the CO₂-sensitive productivity of European airlines from 2000 to 2010

Davide Scotti a,*, Nicola Volta b,1

a University of Bergamo, Department of Management, Information and Production Engineering, via Pasubio 7, 24044 Dalmine, BG, Italy
b Cranfield University, Centre of Air Transport Management, University Way, Cranfield, Bedfordshire MK43 0TR, UK

ABSTRACT

Due to the ongoing increase in the number of commercial flights, greenhouse gas emissions from aviation are expected to rise significantly. Balancing the pursuit of productivity growth with environmental-footprint control policies comprises a long-term regulatory challenge. In this light, the main goals of the present paper are: (i) to measure the CO₂ emissions of European airlines from 2000 to 2010, (ii) to compute airlines’ productivity in developing an environmental-sensitive productivity index, (iii) to compare the obtained results with those resulting from a traditional index, and (iv) to identify the drivers affecting productivity changes. Our results show that on average, airlines’ relative CO₂ emissions have decreased. Although the airlines we studied experienced an average productivity increase—both considering and not considering negative externalities production—environmentally sensible productivity growth is lower than traditional productivity growth. Finally, we find that improvements in load factor as well as a combined increase in stage length and aircraft size affect productivity changes positively, while fuel efficiency is significant only in the case of a CO₂-sensitive measure of productivity.

Introduction

The last decade has been characterised by a globally expanded demand for air transport services, with European airlines exhibiting an increase in available seat kilometres (ASK) and revenue passenger kilometres (RPK) of 18.4% and 26.7% for the 2000–2010 period.2 However, global events such as the 9/11 terrorist attacks, the outbreak of severe acute respiratory syndrome (SARS) in 2003, and the 2007 financial crisis affected European (and global) airlines’ business. These shocks impacted airlines’ performances and reduced the demand for air transport services. Besides the previously mentioned events and the changes in the structure of the European airlines (e.g., mergers and/or reorganization of some of the major carriers), the last decade has also been distinguished by an increasing interest in the environmental impact of aviation. The air transport industry is a growing source of greenhouse gases (GHG), especially carbon dioxide (CO₂) emissions. Under a high-growth scenario, projected 2050 global CO₂ aviation emissions are estimated to be 7–8 times higher than 1990 levels.3 It is clear that the challenge is to create an air transport system that is both efficient and sustainable. Increased awareness of environmental emissions has...
promoted policymakers to adopt measures aimed at considering the environmental impact in economic polices. For example, the EU stated the inclusion of aviation in its Emission Trading Scheme (ETS), which was expected to start in January of 2012, but was postponed after intense criticism (Girardet and Spinler, 2013). Obviously, the objective of a mere increase in productivity without considering undesirable externalities is inconsistent with this attitude. However, traditional measures of productivity growth concentrate only on the production of desirable outputs and fail to consider environmentally negative outputs. This paper is intended as a contribution by offering the following objectives: (i) to assess the CO\(_2\) emissions of 18 major Association of European Airlines (AEA) airlines for the 2000–2010 period, (ii) to develop a productivity index that takes into account production of both positive outputs and bad externalities, (iii) to compare the results of the environmental-sensitive productivity index with a traditional, that is, not environmental-sensitive measure, and (iv) to investigate the main determinants of airlines’ productivity changes.

**Methodology**

The Malmquist index is the most frequently used indicator to compute the productivity of industries over time periods and the most popular approach was proposed by Färe et al. (1995). The peculiarities of this method are that (i) non-parametric frontier technologies are used in computing the index and, (ii) it is possible to decompose the index into technical efficiency and technology change. Moreover, the indirect advantage of this approach is that in using non-parametric technologies there is no need to impose an a priori functional form on the production function. The Malmquist index is widely applied in air transportation to compute airline productivity. Recent examples of applications can be found in Sickles et al. (2002), Greer (2008), Chow (2010), Assaf (2011), and Pires and Fernandes (2012). However, the Malmquist index approach as proposed by Färe et al. (1995) received criticism for a lack of feasibility under certain conditions (Ray and Desli, 1997). Similar drawbacks are present in the Malmquist–Luenberger index, developed by Chung et al. (1997) to consider negative externalities in productivity evaluations. To overcome these problems, we develop a biennial Malmquist–Luenberger index (BML) starting from the biennial Malmquist index (BM) presented by Pastor et al. (2011). The adoption of a “biennial” approach solves (i) infeasibility under variable returns to scale, (ii) the non-identification of technical regress, and (iii) the recalibration of indices when adding new time periods.

We start by representing the production possibility set for \(k = 1, \ldots, K\) airlines producing \(M\) desirable outputs, \(y = 1, \ldots, M\) with \(y \in \mathbb{R}^M\), \(J\) undesirable outputs, \(b = 1, \ldots, J\) with \(b \in \mathbb{R}^J\) by using \(N\) inputs, \(x = 1, \ldots, N\) with \(x \in \mathbb{R}^N\), as follows:

\[
P(x) = \{(y, b) | x \text{ can produce } (y, b)\} \tag{1}\]

Following Färe et al. (2007), our technology is required to satisfy the standard axioms of (i) inactivity, (ii) finite production and consumption levels, and (iii) free input disposability. Regarding the undesirable output production, the technology requires (iv) weak disposability of undesirable outputs, (v) strong disposability of good outputs, and (vi) outputs null-jointness. The distance from the observations to the frontier is computed by a directional distance function (DDF).\(^4\) Our DDF seeks to increase the desirable outputs while simultaneously decreasing the undesirable outputs. Formally, it can be defined as follows:

\[
\bar{D}(x, y; b; g_y, g_b) = \max \{\beta | (y + \beta g_y, b - \beta g_b) \in P(x)\}, \tag{2}
\]

where \(g = (g_y, g_b)\) is a direction vector with \(g \in \mathbb{R}^M \times \mathbb{R}^J\), defining the direction of the increase (decrease) of desirable (undesirable) outputs. Following Pastor and Lovell (2005) and Pastor et al. (2011), we define a series of overlapping biennial technologies for each pair-wise comparison of adjacent periods. Two adjacent time periods are sufficient to overcome the problems normally affecting the previous productivity indices.

We compute the defined technologies and respective distance functions using a variable return to a scale data envelopment analysis approach (DEA). Starting from the biennial technology and following Chung et al. (1997), we can define the Biennial Malmquist–Luenberger index (BML) as follows:

\[
BML(x_t, y_t, b_t, x_{t+1}, y_{t+1}, b_{t+1}) = \frac{1 + \bar{D}^b(x_t, y_t, b_t)}{1 + \bar{D}^b(x_{t+1}, y_{t+1}, b_{t+1})}, \tag{3}
\]

where the directional distance function \(\bar{D}^b(x, y, b)\) is defined on the biennial benchmark technology. A BML index \(> (\leq 1\) means that the production enables more (less) desirable outputs and less (more) undesirable outputs for a given amount of inputs. Consequently, a BML = 1 implies the absence of productivity growth. Moreover, since the technology includes both period \(t\) and period \(t + 1\) technologies, we do not need to apply the usual geometric mean in defining Eq. (3). The BML index can be decomposed into the components of productivity growth as follows:

\[
BML = ECxTC, \tag{4}
\]

\(^4\) As stated in Aparicio et al. (2013), the DDF approach may result in a possible inconsistency problem in the technical-change computation. However, the solution proposed causes the violation of a null-jointness axiom and creates a disputable technology design (Arabi et al. 2015). We therefore do not implement the Aparicio et al. correction since the inconsistency problem could appear in just a few specific situations.
where $EC$ is the efficiency change and $TC$ is the technical change. The former represents the efficiency gain over the time period and the latter accounts for technical progress in the airlines’ production function. $EC > (<) 1$ means an increase (decrease) in the efficiency of the decision-making unit (DMU) with respect to the contemporaneous frontiers. $TC > (<) 1$ indicates that the best practice technology in the period $t + 1$ is closer (further) to the biennial best practice frontier than the best practice technology in the period $t$. In other words, when the TC is greater (smaller) than one, there is a progress (regress) in the technology. To calculate and decompose the BML index, we need to specify four DDFs: two functions with observations under consideration and the biennial production set—namely, $D^B_t(x^t, y^t, b^t; g^t, s^t)$ and $D^B_{t+1}(x^{t+1}, y^{t+1}, b^{t+1}; g^{t+1}, s^{t+1})$—and two functions in which observations and the production set are from the time periods under observation, that is, $D^I_t(x^t, y^t, b^t; g^t, s^t)$ and $D^I_{t+1}(x^{t+1}, y^{t+1}, b^{t+1}; g^{t+1}, s^{t+1})$. Following Chung et al. (1997) and Kumar (2006), for a BM and BML comparison, we consider the direction vector $g = (y, -b)$. Thus, we seek the maximum proportional contraction of negative externalities and the maximum proportional expansion of positive outputs.

Finally, the biennial Malquist (BM) index is computed as presented in Pastor et al. (2011). The main difference between the BM and BML indices is that the BM does not consider the production of negative externalities in the distance computations. Oh (2010) describes the equivalence between the global Malquist–Luenberger and the global Malquist indices and such equivalence is totally valid for the BM and the BM index.

**Airlines’ production process**

In the airlines business, it is crucial to exploit the available capacity since operating flights with empty seats (or empty cargo space) is not only extremely inefficient, but also costly from an economic point of view. We benchmark airlines based on their ability in maximising capacity utilisation rates. Total capacity offered by airlines can be divided into two components: (i) the number of available seat kilometres (ASK) as a measure of capacity reserved to passenger service and (ii) the available freight ton kilometres (AFTK) as a measure of cargo capacity. These two input variables depend on airlines’ strategies in terms of combined aircraft sizes, flight frequencies, and routes offered. Airlines can exploit this capacity by transporting passengers and/or freight. Hence, the desirable production-process outputs include (i) the number of revenue passenger kilometres (RPK), and (ii) the revenue tonnes of freight kilometres (TFTK). Finally, the CO2 emitted represents an undesirable output of airlines’ production process. This is a simple but effective representation of airlines’ production process. Airlines are efficient when they carry as many passengers and freight as possible and produce as little CO2 as possible given the flights operated (i.e., the capacities inputs). Notice that given the level of inputs, airlines can increase their efficiency (i) by reducing CO2 emissions, which requires operating greener aircrafts, and (ii) by increasing passenger and freight volumes. Moreover, according to our variable selection, we indirectly compare airlines in terms of load factors (i.e., passenger and cargo load factor) and the unitary amount of CO2 produced ($CO2/ASK$ and $CO2/AFTK$ are the bad output/input ratios). Given the objectives of this paper, we do not consider ticket pricing, quality of the service, or airlines’ profitability. However, in a certain sense, the effect of quality of service and fare levels is reflected in airlines’ load factors. In the extant literature, the most used output variables are passenger and freight volumes (Barla and Perelman, 1989; Cornwell et al., 1990; Oum and Yu, 1995; Baltagi et al., 1995; Coelli et al., 1999; Barbot et al., 2008) usually measured in RPK and total freight ton kilometres (TFTK). On the input side, there is variety with the most popular variables represented by capital assets (e.g., capital expenses and number of planes), capacity (e.g., aircraft capacity, available seat kilometres) and operational costs (e.g., fuel and/or labour costs/quantities). If our output choice is consistent with the previous research, our input variables slightly differ since labour and fuel are not considered in the analysis. The motivation for our choice is that fuel consumption is directly proportional to CO2 production. This proportionality could cause biases in the DEA models used to compute the productivity indices. Moreover, data envelopment analysis in order to produce a meaningful production frontier requires substitutable inputs and complementary outputs (Bogetoft and Otto, 2011). The common choice of using capacity and labour as inputs in a DEA approach may raise concerns about the substitutability requirement (e.g., moving on input isoquants, it is possible to use more (less) labour and less (more) capacity to carry the same amount of passengers). Our variable selection meets the requirements of substitutability and complementarity while effectively describing the airline production process.

**The dataset**

Our sample includes 18 major European airlines members of the Association of European Airlines (AEA) over the 2000–2010 period. We consider airlines belonging to European Union state member countries according to the European regions defined by ISO/IATA. We exclude from the analysis carriers not operating every single year (e.g., because of failure) as well as low-cost carriers (e.g., Ryanair, EasyJet, etc.), and purely cargo carriers (e.g., Cargolux, DHL, etc.) because of their different business models. Low-cost carriers do not usually operate cargo or belly flights, and conversely, cargo carriers do not operate passenger flights. Considering those carriers would lead to units producing zero outputs and consequently to biased DEA results.

Airlines are included with respect to the operating International Air Transport Association (IATA) code. In cases of mergers during the period considered, if the airline maintained an IATA code, it is considered an individual observation (e.g., Air
France and KLM after the 2004 merger, Austrian Airlines after the acquisition by Lufthansa in 2009). Table 1 shows the carriers we include in the analysis and their relative IATA codes.

For each flight operated by the selected airlines during the analysed period, we collect the following data from the Official Airline Guide (OAG) database:

- scheduled available seats (AS);
- scheduled available seats kilometres (ASK);
- scheduled available freight capacity in tonnes (AFT). Freight capacity in belly hold, combined flights, and purely cargo flights are included but airlines’ freight capacity on other transport modes are filtered and not considered;
- great-circle distance;
- aircraft model.

By aggregating the single origin–destination flights we obtain yearly values for AS, ASK, available freight tonne-kilometres (AFTK) and theoretical kilometres flown for each airline. OAG data consider scheduled flights and default aircraft model capacities and our data may differ slightly from the real values. Regarding non-scheduled flights, the underestimation seems to be irrelevant given the magnitude of the values (e.g., in 2010, the Lufthansa non-scheduled ASK counted for the 0.09% of the total ASK, source: ICAOdata+). Regarding the default aircraft settings, we are aware that airlines could modify capacity allocation because of payload considerations (e.g., more passengers and less cargo) resulting in different available seats/freight ton. Nevertheless, real capacity data for each airline as well as aircraft and route combination were not available. Finally, the use of OAG data allows for a detailed estimation of CO₂ production since information on aircraft models and great-circle distances are available for each specific flight.

The production outputs are collected from the AEA Monthly Traffic Update. We considered the scheduled revenue passenger kilometres (RPK, millions), and the scheduled total freight tonne-kilometres (TFTK, thousands). The output data are selected to be consistent with the inputs OAG data. In the few cases of a lack of information or correspondence among the different data sources, we integrate the data using the airlines annual reports.

In order to compute the CO₂ emissions we adopt the following procedure:

1. We collect the airline, the aircraft model and the distance of the operating flights from the OAG database. By considering operating flights we are able to exclude non-operating duplicate flights arising from code share agreements. We use the great-circle distances since data on real flown distances were not available. We note that the use of great-circle distances may result in an underestimation of CO₂ production.
2. We use the Eurocontrol emission calculator (2012 version) to obtain the CO₂ produced by each couple aircraft model/distance. The calculator estimates the CO₂ produced both during the flight and the landing take-off cycle for each specific aircraft model. Unfortunately, the tool (and OAG) does not differentiate between engine models installed, thus reducing the estimation accuracy.
3. We aggregate the CO₂ produced across all flights in order to obtain the total amount of CO₂ emitted yearly by each airline.

Table 2 includes the descriptive statistics of the variables considered in the productivity analysis. Our sample covers about 90% of the 2010 total ASK provided by AEA members and 60% of the 2010 European airlines ASK according to the OAG database.

**European airlines’ CO₂ production levels**

Fig. 1 shows the amount of CO₂ kilograms per available seat kilometres (ASK) and per revenue passenger kilometre (RPK) of each European airline included in our sample for years 2000 (represented with black points) and 2010 (grey points). The CO₂/ASK is a value strongly influenced by the aircraft model owned by the airline (greener aircraft means, ceteris paribus, a lower level of CO₂ emissions), but also by the network offered (longer flights may result, ceteris paribus, in lower CO₂/ASK values given that the CO₂ produced during the landing and take-off cycle is divided by a larger amount of kilometres). The CO₂/RPK is additionally influenced by the airlines’ capability of filling the seats offered. In this sense, a lower distance
between the two ratios is a signal of a higher average airline’s load factor. Fig. 1 does not consider the CO2 produced by purely cargo flights in order to fairly compare airlines.

Dotted lines represent the average values (grey for year 2010 and black for 2000). The general shift of the observations towards the bottom-left side of the picture reveals a decrease over time in both CO2/ASK and CO2/RPK. Average values decrease over 10 years of 6% (from 0.101 to 0.095 of CO2/ASK) and of 11% (from 0.143 to 0.127 of CO2/RPK). However, not all the airlines experienced this evolution—a few exhibited increases either in CO2/ASK (Adria Airlines (JP), Virgin Atlantic (VS), British Airways (BA) and Cyprus Airways (CY)), or in CO2/RPK (Air Malta (KM)). The relevant distance between the horizontal lines (i.e., the sample averages of CO2/ASK) seems to cross-refer a fleet evolution towards more modern aircraft. In this regard, Fig. 2 lists the aircraft families’ utilisation levels for 2000 (black) and compares them with the 2010 (grey) levels in terms of percentage of movements (2a) and ASK (2b). Fig. 2(a) highlights the changes in narrow-body aircraft utilisations since this type of aircraft is used in high-frequency, short-medium distance routes. Differently, considering the ASK (2b) we analyse the wide-body aircraft, high-capacity aircraft used mainly on long distance routes. The evidence suggests that for both narrow and wide bodies, airlines have replaced fuel-inefficient aircraft models. As an example, the MD-80 utilisation rate decreased from 14.4% of total flights (year 2000) to only 2% (year 2010) while the A320 utilisation rates increased from 14.2% in 2000 to 36.6% in 2010. The estimated fuel consumption of the MD-80 is around 4,600 kg with an estimated emission of 14,500 kg of CO2 for a 1,000 km flight while for the same distance the A320 consumes 4,000 kg of fuel emitting 12,600 kg of CO2.

5 The same trend is observed for wide-body aircraft on long-haul routes. The relevant increase in the utilisation rate of Boeing 777 aircraft in terms of ASK (from 6.6% to 18.1%) and the decrease of the use of Boeing 747 (from 40.5% to 19.1%) highlight the global switch to more fuel efficient and less environmental-impacting aircraft as a partial reaction to the increases in fuel costs.

Traffic data

An initial insight into the sample level concerning airlines’ productivity over time can be seen in Fig. 3. Specifically, Fig. 3(a) shows the yearly change (with respect to the base case year 2000) of both the offered passenger capacity—measured

---

Table 2
Descriptive Statistics of Input and Output Variables.

| INPUTS | DESIRABLE OUTPUTS | UNDESIRABLE OUTPUT |
|--------|-------------------|---------------------|
| ASK (.000,000) | AFTK (.000,000) | RPK (.000,000) | TFK (.000,000) | CO2 ton (.000) |
| Mean | 46,199 | 2,774 | 33,783 | 1,499 | 4,842 |
| Median | 23,987 | 818 | 16,202 | 360 | 2,093 |
| Min | 1,005 | 14 | 678 | 2 | 97 |
| Max | 181,281 | 13,621 | 131,669 | 8,346 | 20,522 |
| Std. Dev. | 53,146 | 3,844 | 39,158 | 2,236 | 5,833 |

Fig. 1. European airlines’ average levels of CO2/ASK and CO2/RPK (2000 and 2010).

5 Source: Eurocontrol emission calculator. The values are rounded averages for purely comparative purposes.
in terms of available seat kilometres—and the passenger volume—measured in terms of revenue passenger kilometres. The measures increase from 2002 to 2007 when the first effects of the financial crisis are recognisable in a slowdown of the RPK increase. From 2008 to 2009, both ASK and RPK present a fall-down given the quickening of the global crisis. Since 2004, the RPK (dotted) line remains consistently above the ASK line (solid line), meaning that passenger traffic has increased at a
higher rate than passenger capacity. This demonstrates airlines’ great effort in avoiding spare capacity as well as close attention paid to passenger load factors. Fig. 3(b) shows the same comparison in terms of cargo. The two lines often intersect over the period analysed suggesting an irregular path for the cargo load factor. Interestingly, the solid line (capacity) remains consistently above the dashed line (carried volumes) from 2005, highlighting that airlines overestimated the increase in cargo demand and offered more capacity than needed. These insights are confirmed in Fig. 3(c), in which passengers and cargo load factor (computed as RPK/ASK and TFTK/AFTK) are compared with their base cases (year 2000). Finally, Fig. 3(d) displays the trend in terms of CO₂ produced per ASK and per RPK. These measures show the pollution produced per capacity and traffic units. The average decreasing trend confirms the idea inferred from Figs. 1 and 2: airlines’ eco-efficiency seems to increase over time due to technological progress—i.e., a greener and more fuel-efficient aircraft fleet. Fig. 3(d) does not consider emissions due to cargo flights. Nevertheless, passenger traffic (i.e., the number of flights) is predominant and can be considered a valid proxy of the whole situation.

Fig. 4 shows additional information that completes the picture of airlines’ evolution over time. First, networks seem to evolve towards longer flight distances (the average flight lengths increase from 1,198 km in 2000 to 1,407 km in 2010). Second, the average decrease in fuel consumption—from 6.86 kg/km (year 2000) to 6.27 kg/km (year 2010)—may suggest that airlines did improve fuel efficiency. Last, in analysing aircraft size of passenger flights, after an initial period (2000–2002) of size decrease and a second period (2002–2006) of a near-constant trend, from 2007 onwards, airlines seem to operate bigger aircraft given the increase in the average available seats per flight (from 122 seats in 2006 to 132 in 2010—i.e., +8.2% in average aircraft size). This may partially reflect a reaction to higher airport congestion and increased slot values that force airlines to adopt larger aircraft. However, other practices may influence the per flight average available seats such as (i) an increase/decrease of long-haul flights or (ii) a change in seat configurations (e.g., reducing first class/business seats in favour of smaller economy seats). A theoretical discussion consistent with our empirical sample evolution can be found in the research by Brueckner and Zhang (2010). Among other results, the authors highlight how the increase in emission charges (or fuel prices in Europe during the decade analysed) leads to an average increase in load factors and stage length.

Figs. 3 and 4 draw only partial pictures of the situation and additional information is required to study airlines’ total productivity changes. In order to capture both technical and efficiency improvements over time, an evaluation of productive performances by proper instruments is necessary, such as the indices presented in Section “Methodology”.

Results and discussion

BML and BM trends

Following the approach presented in Section “Methodology”, we compute the best practice biennial technical frontier from our sample. By decomposing the index into its components, we are able to analyse different sources of productivity change, namely catch-up effects (efficiency change) and technical change. BML (solid line) and BM (dotted line) indices
are presented as unweighted average rates of change in Fig. 5. Hence, it is possible to observe the performance of the average European airline recalling that values greater (lower) than the unity imply an improvement (deterioration) of the relevant index.

Concerning the productivity change, our model suggests an increase, on average, of productivity from 2001 to 2007 and from 2009 to 2010 in both the BML and the BM index. The average rate of the BM productivity shows a slowdown in 2002 (from 3.48% in 2001 to 1.80% in 2002) probably due to the effect of the 9/11 terrorist attacks and a regress of productivity at the beginning of the financial crisis in 2008 (−2.21%). Similarly, the BML average rate of productivity shows a slowdown in 2002 and a unique small regress during the financial crisis (−0.65%). The 2008 regress of BM and BML productivities are ascribed to an efficiency and technical deterioration (Fig. 5(b) and (c)). The 2003 SARS pandemic does not appear to affect the productivity and efficiency change although it does impact the technical change indices. The EC index captures the distance of the observations to their respective best practice frontiers, while an increase in the TC index registers an expansion in the PPS. Such effects, marked in the BM indices, are muffled by the inclusion of bad externalities in the BML indices. Similar results can be observed in Oh (2010). Our results suggest that during the 2003 SARS pandemic, airlines were successful in planning a capacity reduction (i.e., contracting the production possibility set) to anticipate the decrease in demand and maintaining a high level of load factor (i.e., increasing the EC). With regards to the economic crisis, our results suggest that despite a contraction of the production possibility set (i.e., a reduction in capacity offered) the demand fall was worse than expected and led to a decrease in airlines’ efficiency (i.e., lower load factors). The results obtained are coherent with the general trends shown in Fig. 3, especially those referring to the 2008 shortfall. Indeed, in comparing years 2007 and 2008, a decrease is evident in both passenger and cargo load factors due to the slowdown in output levels (i.e., RPK and TFK) that were not compensated by a proportional decrease in input levels (i.e., ASK and AFTK). Interestingly, the shortfall is more marked in the BM productivity index, which may be due to the fact that the BML index incorporates the positive effect of the decrease in CO₂ per ASK between 2007 and 2008 (Fig. 3(d)). Moreover, the increase in productivity for both BM and BML indices in 2009 are reasonable and in line with Fig. 3. Despite a general decrease in passenger and cargo traffic, the lower level of capacity offered (Fig. 3(a) and (b)) reflects airlines’ reaction to the crisis resulting in a higher average passenger load factor (Fig. 3(d)).

As in the majority of empirical studies analysing productivity changes, Fig. 5 shows the unweighted mean for indices. Nevertheless, it is of interest to account for the relative importance of each observation in order to obtain a representative productivity value at an industry level. Different solutions were developed in the literature to weigh productivity indices. Zelenyuk (2006) develops a Malmquist index aggregation based on prices, while Coelli and Rao (2005), among others, use the DEA output values as weights. Unfortunately, the former methodology cannot be implemented since price information is not available for our application. Regarding the use of the output as weight, we are facing a conceptual problem since we apply a multi-output approach (i.e., for the BM index our outputs are RPK and TFK, while for the BML we consider RPK, TFK and
In the case of the BM index, we can use as weight the aggregation of the two traffic outputs (in our case, we use the work load units, WLU), while an aggregation of traffic data and CO$_2$ production is not possible for the BML. Thus, we consider the WLU as weight for both the indices. Using a unique traffic-aggregation weight could produce biases when comparing the weighted BM and BML since only part of the airlines output is considered for the latter index. Even without a comparison between the weighted indices, it remains of interest to analyse the productivity changes of the European airlines industry when accounting for the output level (i.e., the relative size of the airlines). Fig. 6 shows the two weighted BM and BML indices. Three main observations can be drawn by comparing Figs. 5(a) and 6: (i) in 2001 the weighted indices capture a stronger effect of 9/11 by highlighting a productivity regress, (ii) the 2003 SARS epidemic shows a bigger impact, and (iii) productivity indicates a delayed and longer regress due to the financial crisis. Clearly, weighting for the work-load units increase the importance of bigger airlines, which are naturally more exposed to global events. It turns that 9/11, SARS, and the financial crisis significantly impact productivity changes.

### Airlines’ heterogeneity

The average yearly productivity, efficiency, and technical change computed at the airline level are listed in Table 3. We recall that values greater (less) than unity indicate an improvement (deterioration) in the relevant performance. Three airlines show an average productivity deterioration according to the BML index (bmi, Cyprus Airways, Air Malta) and four according to the BM index (British Airways, bmi, Cyprus Airways, Air Malta). Looking at the average efficiency change, the BML index shows three airlines having efficiency regress (British Airways, bmi, Cyprus Airways), while five airlines experience a decrease in technical efficiency according to the BM index (British Airways, bmi, Cyprus Airways, Iberia, Air Malta). Finally, concerning the technical change, it is possible to observe two airlines regressing the technology (Cyprus Airways and Air Malta for the BML, while bmi and Air Malta for the BM).

---

7 Work load units (WLU) are an aggregate measure of passenger and cargo volumes with 100 kg of freights equalised to one passenger.
As indicated in the last row of Table 3, the yearly BML productivity growth (1%) is, on average, lower than the BM growth (1.2%). In other words, the environmental-sensitive index shows a lower yearly rate of increase with respect to the typical productivity growth index. This may suggest that, on average, a decrease in the level of CO₂ emissions is slower or less than the increase in RPK and TFTP. In fact, in the short-medium term, it is easier for airlines to increase the RPK and TFTP rather than reduce CO₂ emissions. Ceteris paribus, there are only two ways to reduce CO₂ emissions: (i) to substitute current aircraft with greener ones or (ii) to upgrade part of the aircraft (e.g., engines or engine components). The former solution is more effective but costly while the latter is cheaper but less effective. Generally, aircraft lifecycle is around 20 years and not all aircraft in a fleet are substituted at the same time. Thus, given the time frame analysed, our results are catching only partial fleet changes (as shown in Fig. 2). Nevertheless, some airlines exhibit a BML score greater than the BM scores. As an example, British Airways obtains an average BML growth while registering an average productivity decrease according to the BM index. This may suggest that including negative externalities in the analysis not only avoids a productivity overestimation, but also provides a more complete measure of airlines’ performance.

Innovative airlines

A value of technical change greater than unity does not necessarily imply that the considered airline pushes the production possibility set (PPS) outwards. In order to state which airline shifts the frontier, a more rigorous analysis is necessary. Several reasons could explain an expansion of the airlines’ PPS. Successful mergers, acquisition of new airlines, new alliance membership, network changes, adoption of new generation aircraft, or winning marketing strategies are all situations in which airlines could be innovators. The airlines industry is characterised by high competition and complex interactions between technological, political, and economic factors. In the computation of innovative airlines, it is impossible to specifically identify and measure those peculiarities and only a general picture of the industry can be captured. Nevertheless, this approach enables an understanding of the differences and how the PPS reacts when negative externalities are taken into consideration. Methodologically, Färe et al. (2001) defined the set of conditions to determine the innovator units as follows:

\[ \mathcal{T}^{t+1} > 1 \]  \hspace{1cm} (5a)

\[ \mathcal{D}^{t} (x^{t+1}, y^{t+1}, b^{t+1}) < 0 \]  \hspace{1cm} (5b)

\[ \mathcal{D}^{t+1} (x^{t+1}, y^{t+1}, b^{t+1}) = 0. \]  \hspace{1cm} (5c)

The first condition (5a) indicates that the technical frontier shifts in the more good outputs and in lower bad output direction. This means that in period \( t + 1 \), it is possible to increase the RPK and TFTP and reduce the production of CO₂ with respect to period \( t \). The second condition (5b) indicates that production in \( t + 1 \) occurs outside the PPS defined in period \( t \) with the input level of \( t + 1 \), meaning that technical change has occurred during the transition period. The last condition (5c) states that the airline is on its PPS in the period of \( t + 1 \). Table 4 lists the innovative airlines for a consecutive two-year period from 2000 to 2010.

The first consideration is that, for both indices, no airlines are deemed as innovators for the entire study period. Austrian Airlines, Air France, and Czech Airlines appear often as innovators with a BML approach whereas Air France and SAS are innovators in a BM approach. Generally, airlines that are innovators in both indices during the same period were able to balance a productivity growth index. This may suggest that including negative externalities in the analysis not only avoids a productivity overestimation, but also provides a more complete measure of airlines’ performance.

Drivers of productivity changes

Following Zaim (2004) and Yörük and Zaim (2005), we perform a second stage analysis on BM and BML indices. That is, we investigate the impact of a set of airlines’ specifics to identify the main drivers of productivity change. We study whether productivity changes are related to shifts in (i) network characteristics, (ii) airlines’ performance, and (iii) airlines’ fuel efficiency. As network characteristics, we consider the average flight distance (“avg_length”—computed as total kilometres flown divided by the number of flights), the total distance flown (“tot_length”), the average size of the aircraft (“avg_size”—computed as total work load units’ capacity divided by the number of flights), and the total flights (“tot_flight”). Distance flown, aircraft size, and number of flights are three connected decision variables. First, an increase (decrease) in the average flight distance may be due to either a greater (lower) frequency in longer flights or to the introduction of new longer (shorter) routes. Second, operating on longer (shorter) distances may be connected to the utilisation of bigger (smaller) aircraft. Third, operating smaller aircraft requires a higher number of flights to carry the same amount of passengers and freight. As far as airlines’ performances, we include both passengers and cargo load factors (“passenger_LF” and “cargo_LF”) in the regression. These variables are expected to positively affect the productivity change. Finally, to investigate the influence of fuel efficiency on productivity, we consider the average fuel per kilometre burn as a level of the fleet’s

---

8 The same conditions are valid for the BM index in which only good outputs are considered.
ecological impact (“fuel_km”—computed as the total amount of fuel consumed divided by the total distance flown). While it is not clear a priori the effect of fuel efficiency on the BM index, we expect it to affect positively the BML due to the relationship between fuel burnt and CO2.

Given that the BML index (as well as the BM index) evaluates the shift in productivity between two consecutive years, the determinants are introduced as variations between two periods (\( t/C_0 \)). Each variation is expressed as the ratio between the value at time \( t \) and the value at time \( t/C_0 \). As a result, all the variables are greater (lower) than unity consistent with the dependent variable. Let the biennal Malmquist index \( BML_i \) represent the productivity change from period \( t/C_0 \) to period \( t \) for airline \( i \). Eq. (6) specifies the relation between the BML index and its determinants as follows:

\[
BML_i = \beta_0 + \beta_1avg\_length_i + \beta_2avg\_sizei + \beta_3(avg\_length_i * avg\_sizei) + \beta_4tot\_flight_i + \beta_5tot\_length_i + \beta_6passenger\_LF_i + \beta_7cargo\_LF_i + \beta_8fuel\_km_i + \epsilon_i, \quad (6)
\]

where \( i \) is the airline and \( \epsilon \) is the disturbance term. The same equation is applied with \( BM_i \) as the dependent variable. Estimations are based on a pooled OLS regression technique and the results are shown in Table 5 (BML in the second column and BM in the last column).

Starting from the BM index, in order to analyse the effects of the basic terms (\( avg\_length \) and \( avg\_sizei \)), the marginal effects must be computed. The partial derivatives show that an increase of the stage length has a negative impact on productivity while the basic impact of the \( avg\_sizei \) is not significant. Analysing the interaction term, airlines are likely to boost the productivity combining the increase in the average aircraft size and in the stage length. In other words, only using bigger aircraft on longer routes could increase the productivity, while a non-connected increase seems to be deleterious on the index. This evidence is reinforced by the positive impact of total kilometres and the negative impact of the number of flights. These results may be interpreted again as a signal that in order to increase productivity, an airline has to generally decrease movements and serve longer routes only if using bigger aircraft. These findings are in line with recent scenarios of full-service carriers

---

Table 4

| Period       | BML innovative airlines | BM innovative airlines |
|--------------|-------------------------|------------------------|
| 2000–2001    | OS – Austrian           | SK – SAS Scandinavian Airline |
| 2001–2002    | AF – Air France         | AF – Air France        |
|              | KL – KLM                | KL – KLM               |
|              | OK – Czech Airlines     | OK – Czech Airlines    |
|              | TP – TAP Portugal       |                         |
| 2003–2004    | CY – Cyprus Airways     | KL – KLM               |
|              | KM – Air Malta          | KM – Air Malta          |
|              | LH – Lufthansa          | LH – Lufthansa          |
|              | TP – TAP Portugal       | TP – TAP Portugal       |
| 2004–2005    | CY – Cyprus Airways     | CY – Cyprus Airways     |
|              | OS – Austrian           | SK – SAS Scandinavian Airline |
| 2005–2006    | AF – Air France         | AF – Air France        |
|              | IB – Iberia             | RO – TAROM             |
|              | OK – Czech Airlines     |                         |
| 2006–2007    | AF – Air France         | AF – Air France        |
|              | IB – Iberia             | SK – SAS Scandinavian Airline |
|              | VS – Virgin Atlantic Airways | VS – Virgin Atlantic Airways |
| 2007–2008    | KM – Air Malta          | MA – Malev Hungarian Airlines |
|              | MA – Malev Hungarian Airlines |                     |
| 2008–2009    | AZ – Alitalia           | AZ – Alitalia           |
| 2009–2010    | AY – Finnair            | AZ – Alitalia           |
|              | AZ – Alitalia           |                         |
|              | OK – Czech Airlines     | VS – Virgin Atlantic Airways |
|              | TP – TAP Portugal       |                         |
|              | VS – Virgin Atlantic Airways |                     |

---

9 Various specification tests performed reveal the absence of multicollinearity and homoschedastic, independent, and approximately normal distributed residuals.

10 The partial derivatives of \( avg\_length \) are equal respectively to 2.642 in the BML regression and 3.633 in the BM regression (both significant at the 1% level), while the partial derivatives of \( avg\_sizei \) are equal respectively to 0.089 in the BML regression and 0.036 in the BM regression (both not significant).
focusing on international and intercontinental flights to decrease low-cost carriers’ competitive pressure. As expected, an increase in both passengers and cargo load factors positively affects airlines’ productivity. Finally, a change in the fuel efficiency does not seem to significantly influence the productivity index. When CO2 is included in the analysis (BML index), we can observe the same conclusion with respect to the network characteristics and the airline performances. However, a reduction in the fuel per km consumption becomes a significant determinant of productivity change due to the fact that lower amounts of fuel burnt produce lower CO2 emissions, ceteris paribus. This result enforces the evidence that operating greener aircraft is the primary way to increase the environmental productivity.

**Conclusion**

A biennial Malmquist–Luenberger index was developed to determine CO2 production in airlines’ productivity assessment. We consider an updated modelling approach to compute a correct measure of productivity, technical, and efficiency change. We compare an environmental-sensitive index (BML) and a traditional index (BM) that considered a sample of 18 major European traditional airlines for the 2000–2010 period. The airlines’ production process is analysed on the basis of two input variables describing capacity—carriers’ capacity in terms of passengers and cargo—and three output variables, two desirable describing the perceived traffic levels (i.e., passengers and cargo volumes) and one undesirable describing the environmental impact in terms of CO2 emissions. Thus, we benchmark airlines’ performances in exploiting the amount of capacity generated by network structure and the fleet mix/flight frequency combination. Finally, we regress the two indexes on three sets of variables (network characteristics, airlines’ performance, and fuel efficiency) to evaluate the main drivers of productivity changes.

Our results show that both the BML and BM indices present shortfalls corresponding to the financial crisis years. In addition, the environmental-sensitive index presents a lower rate of increase with respect to the typical productivity growth index. Thus, on average, despite a general decrease in CO2 emissions, we highlight that airlines emphasize the increase in RPK and TFTK more. In this sense, we also show that including negative externalities avoids productivity change overestimations and provides performance measurements in line with the need for limiting emissions and climate change. Furthermore, we show that improvement in load factors and a combined increase of stage length and aircraft size foster both traditional and CO2 sensitive productivity growth. Finally, fuel efficiency is only significant as a driver of BML productivity growth.

**Acknowledgements**

We would like to thank Prof. Martin Dresner for his help and interest in our work. We also thank the participants of the Air Transport Research Society (ATRS) Conference 2013 for their valuable advice.

**References**

Aparicio, J., Pastor, J.T., Zofío, J.L., 2013. On the inconsistency of the Malmquist–Luenberger index. Euro. J. Oper. Res. 229 (3), 738–742.
Arabi, B., Munisamy, S., Emrouznejad, A., 2015. A new slacks-based measure of Malmquist–Luenberger index in the presence of undesirable outputs. Omega 51, 29–37.
Assaf, A., 2011. A fresh look at the productivity and efficiency changes of UK airlines. Appl. Econ. 43 (17), 2165–2175.
Baltagi, B.H., Griffin, J.M., Rich, D.P., 1995. Airline deregulation: the cost pieces of the puzzle. Int. Econ. Rev. 36 (1), 245–259.
Barbot, C., Costa, A., Sochirca, E., 2008. Airlines performance in the new market context: a comparative productive and efficiency analysis. J. Air Transport Manage. 14, 270–274.
Barla, P., Perelman, S., 1989. Technical efficiency in airlines under regulated and deregulated environment. Ann. Public Cooper. Econ. 60 (1), 103–124.

| Variables     | BML       | BM       |
|---------------|-----------|----------|
| avg_length    | −6.658    | −9.050   |
| avg_size      | −3.987    | −5.460   |
| avg_length avg_size | 4.034    | 5.441    |
| Tot_flight    | −2.398    | −3.294   |
| tot_length    | 2.410     | 3.297    |
| LF_passenger  | 0.243     | 0.444    |
| LF_cargo      | 0.043     | 0.100    |
| fuel_km       | −0.218    | −0.147   |
| Constant      | 7.536     | 9.676    |
| Adjusted R²   | 0.361     | 0.407    |
| Observations  | 180       | 180      |

Values in parenthesis represent the standard errors.
* Significance at 10%.
** Significance at 5%.
*** Significance at 1%.
