Transportation Research Part B 166 (2022) 444–467

Contents lists available at ScienceDirect

Transportation Research Part B

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

# Does big data help answer big questions? The case of airport catchment areas & competition

Nicole Adler [a], Amir Brudner [a,*], Riccardo Gallotti [b], Filippo Privitera [c], José J. Ramasco [d]

[a] Hebrew University Business School, The Hebrew University of Jerusalem, Mount Scopus, 91905 Jerusalem, Israel
[b] Fondazione Bruno Kessler, Via Sommarive 18, 38123, Povo (TN), Italy
[c] Cuebiq Inc., 45 W 27th Street, 3rd floor, 10010, New York, NY, USA
[d] Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), 07122 Palma de Mallorca, Spain

ARTICLE INFO

*Keywords:*
Data analytics
Game theory
Transportation markets
Aviation

ABSTRACT

We develop algorithms to analyze Information and Communication Technologies (ICT) data in order to estimate individuals' mobility at different spatial scales. Specifically, we apply the algorithms to delineate airport catchment areas in the United Kingdom's Greater London region and to estimate ground access trip times from a very large ICT dataset. The spatial demand is regressed over demographic, socio-economic, airport-specific and ground access modal characteristics in order to determine the drivers of airport demand. Drawing on these insights, we develop a catchment area game inspired by Hotelling that analyzes the potential impact of collaboration between airports and airlines by integrating evidence of consumer behavior with producers' financial data. We apply the game to a case study of two London airports with overlapping catchment areas for local residents. Our assessment of airline-airport vertical collusion and airport-airport horizontal collusion indicates that the former is beneficial to both producers and passengers. In contrast, whilst horizontal and vertical collusion is the equilibrium outcome in the analytic symmetric case, it is found to be less likely in the asymmetric case and the real-world, data-driven analysis, due to catchment area and cost asymmetries. Thus, such new datasets may enable regulators to overcome the long-standing information asymmetry issue that has yet to be resolved. Combining new data sources with traditional consumer surveys may provide more informed insights into both consumers' and producers' actions, which determines the need (or lack thereof) for regulatory intervention in aviation markets.

## 1. Introduction

The ever-increasing number of big datasets produced by Information and Communications Technologies (ICT), complemented by advancements in the computational power and methodological tools necessary for their assessment, has opened the way to the study of socio-technical systems at unprecedented detail (Vespignani, 2012). A case in point is the wave of new studies that touch several aspects of human long-range mobility, such as seasonal changes in population distribution (Deville et al., 2014), migration (Simini et al., 2012), tourism (Lenormand et al., 2015a; Bassolas et al., 2016), and air passenger flows (Hawelka et al., 2014). ICT has

* Corresponding author.
*E-mail addresses:* nicole.adler@mail.huji.ac.il (N. Adler), amir.brudner@mail.huji.ac.il (A. Brudner), rgallotti@fbk.eu (R. Gallotti), fprivitera@cuebiq.com (F. Privitera), jramasco@ifisc.uib-csic.es (J.J. Ramasco).

https://doi.org/10.1016/j.trb.2022.10.013
Received 21 July 2021; Received in revised form 5 September 2022; Accepted 30 October 2022
Available online 23 November 2022

been used, for example, to examine the interaction of long-range airline traffic with short-range ground transportation, which has been shown to strongly shape the spreading of epidemics (Balcan et al., 2009). This issue has been investigated by drawing data-informed origin–destination demand matrices from GPS traces of mobile phones (Iqbal et al., 2014), micro-blogging (Lenormand et al., 2014) and location-based social networks (Noulas et al., 2012) (see Barbosa et al. (2018) for a recent review). Improving our understanding of the evolution of the transport system and the ability to optimize demand forecasts, however, necessitates not only access to big datasets but also the formulation of adequate methodological foundations (Hosni and Vulpiani, 2017). Specifically, the economic dimension of transportation ought to be taken into account when attempting demand forecasting (Lenormand et al., 2015b; Lotero et al., 2016a,b; Florez et al., 2018). In the case of airports, given that they serve as a platform that connects airlines with passengers, two separate revenue streams may arise that encourage cross-subsidization such that one price is artificially low and the other too high. For example, low aeronautical charges paid by airlines in order to increase passenger volume may lead to excessively high car parking charges and retail prices at the airport (Thelle et al., 2012). It is, thus, necessary to understand the variables driving airline demand such as pricing (Lenormand et al., 2015b; Lotero et al., 2016b; Florez et al., 2018), ground transport accessibility (Gallotti et al., 2016; Yuan et al., 2012; Gallotti and Barthelemy, 2014, 2015), vehicle accessibility, (Gallotti et al., 2012, 2016), taxi availability (Yuan et al., 2012) and public transport timetables (Gallotti and Barthelemy, 2014, 2015).

We develop a new ICT-based approach in order to shed new light on an important question that has arisen in the literature over the last decade: Do airports compete? This is a significant issue because the answer is likely to impact, among other issues, the need to regulate airport charges. In order to showcase the potential use of large data sources, we have chosen the six airports operating in the Greater London region in the United Kingdom (UK) as a case study. In 1987, British Prime Minister Margaret Thatcher privatized the London airports whilst creating the British Airports Authority, which owned Heathrow, Gatwick and Stansted airports, plus several other airports. In 2008, the UK Competition Commission forced the ownership separation of the London airports in an attempt to encourage greater competition between the facilities. Today, Heathrow, Gatwick and Stansted are owned by separate consortia and companies. Three additional airports currently operate in the Greater London region: London City airport, created in 1987, is privately owned. Luton airport is publicly owned by the local borough council, which, in 1997, issued a 30-year management contract to a public–private partnership. Southend airport is owned by the local council but it has been leased since 1994 to a subsidiary of the Stobart Group, a British infrastructure, aviation and energy company. Since all six airports are now in individual hands, one could assume that there is the potential for competition for airlines and passengers, achieved through pricing- and quality-related strategies. Heathrow also potentially competes for transfer passengers with other international gateways in Europe, such as Paris Charles de Gaulle Airport in France and Frankfurt Airport in Germany.

The question of whether airports in multi-airport regions compete remains an open question in the literature. According to Starkie (2002) and Gillen (2011), the deregulation of the airline industry has afforded airlines the freedom to choose which airports best serve their needs and has increased their willingness to change their network choices accordingly. Furthermore, utilizing the Herfindahl–Hirschman Index and a multinomial logit model, Lieshout et al. (2016) conclude that competition among airports in the United Kingdom is substantial, providing the population with multiple options to access almost all of Europe in less than three hours. However, Pels et al. (2001) argue that airports with a high volume of demand, such as Heathrow, pursue profit maximization whilst remaining a preferred hub, which suggests that even in a multi-airport region, airports are not pure substitutes. Adler and Liebert (2014) find that the purely privatized airports in the UK set higher charges than their public counterparts even when potentially competing with other facilities. Airports may also attempt to differentiate themselves by offering mixed quality services to airlines and passengers as a function of the airline business model.

Ground access mode alternatives and trip times to access and egress airport facilities may also help or hinder competition. Pels et al. (2003) argue that access time is a dominant explanatory factor of airport choice, therefore airport management interested in increasing their market share should actively encourage investments in relatively fast access modes. Similarly, Fuellhart (2007), based on a case study of Harrisburg International Airport in Pennsylvania, argues that access plays an important role in the aggregate decisions of passengers. Given the magnitude of costs required for road and rail infrastructure, accessibility to an airport is usually not a decision made by airport owners[1] but is likely to impact the catchment area and demand levels, which ought to be considered when assessing the likelihood of competition.

Several stylized models have been used to analyze the dynamics of airport competition. Barbot (2009) presents a model to analyze incentives for vertical collusion between airports and airlines. The model is a three-stage game in which airlines choose whether to collude in the first stage. The airports then set their aeronautical charges in the second stage, followed by the airlines setting airfares in the third stage. Barbot finds that in the case of market asymmetry, there are incentives for collusion between airports and airlines. Another model is that of Teraji and Morimoto (2014), which focuses on price competition between airports and their impact on airline network choices. In their two-stage game, after airports set their prices, airlines choose their network configuration, i.e., hub or point-to-point services. They distinguish between social-welfare maximizing and equilibrium network outcomes, showing that under equilibrium, an excessive number of point-to-point networks emerge.

To identify collusive behavior in complex, network-based markets such as the aviation industry, we propose a novel game-theoretic model that describes the strategic behavior between two airport–airline couples. The game structure is, in general, valid for describing competition between providers with non-neutral price structures paid by retailers and consumers (Rochet and Tirole,

[1] One exception is that of Southend airport, owned by the Stobart Group, which built and operates the airport railway station. However, as the group states in the 2020 annual report, the station has reported losses over consecutive periods and the group is planning to sell off the asset https://www.esken.com/media/t3pgplph/stobart-group-limited-2020-annual-report.pdf.

2004). For example, an airport could behave strategically and impact the economic outcome by decreasing airport charges. As a response, airlines might reduce airfares and draw increased passengers to the airport's commercial facilities. Accordingly, the airport may increase commercial revenues over and above the loss in airport charges, thus maximizing airport profits. Moreover, in markets embedded in the physical space such as airports, which physically connect airlines and passengers, the distance to the facility constitutes a cost to the passenger, a consideration that, in addition to the airfares and commercial airport prices, determines the catchment areas for the producers. For this reason, our model of airport competition is based on a catchment area game, which could be applied to other markets such as e-commerce companies, supermarket chains and newspapers.

This data-driven framework represents both methodological and practical contributions to the study of markets, in general, and transportation management, in particular. We develop algorithms that utilize big data from GPS records generated by mobile phones in order to create demand matrices, airport market shares and trip times, all at the statistical area level. We then estimate the importance of the drivers of demand in the market using regression analysis. Our analysis shows that airport size, the ground transport network, public transport services offered and socio-economic characteristics of the statistical areas explain passengers' airport selection in a multi-airport region. We use the insights drawn from these analyses to develop a Hotelling-inspired, catchment area game to estimate the likelihood that airports compete for passengers or horizontally collude to maximize the extraction of consumer surplus. Since airports may act as hubs or serve one major airline customer, we also test whether or not airports choose to collude with their main customer, known as vertical collusion. This formulation is sufficiently general to be relevant to many markets with the potential for both vertical and horizontal collusion. The analytical equilibrium outcome of a two airport–airline pair catchment area game suggests that airport–airline vertical collusion is in the interests of both producers hence, is likely to occur. The analytical outcome also suggests that horizontal airport collusion leads to higher profits, indicating that all parties are likely to collude. However, in the numerical analysis of two airports in the Greater London area that share an overlapping catchment area, we find that the larger airport–airline pair chooses to compete, leading to an equilibria outcome that differs from the symmetric, analytical outcome. In summation, the big data analyses provide detailed information on consumers and the catchment area game provides insight into producers' behavior, which together should aid policy-makers when deciding on the need for airport regulation.

The rest of the paper is organized as follows: In Section 2, we present the methodology and the algorithms developed to integrate and analyze the datasets describing passenger behavior. It continues with a detailed description of the case study and an empirical analysis of the datasets. In Section 3, we explore plausible producer behavior by developing a catchment area game and apply it to the analysis of two London airports. We analyze the game both analytically and numerically. In Section 4, we conclude and discuss possible policy responses and further directions.

## 2. Empirical analysis and consumer behavior

In this section, we present the algorithms developed for extracting and analyzing the ICT data. In Section 2.2, we showcase a case study of the Greater London area, including summary statistics, data validation and then visualize the catchment areas for the six London airports. Subsequently, we present two demand analyses in order to explore the variables affecting consumer behavior, air travel demand and airport choice.

### 2.1. Algorithms to extract data

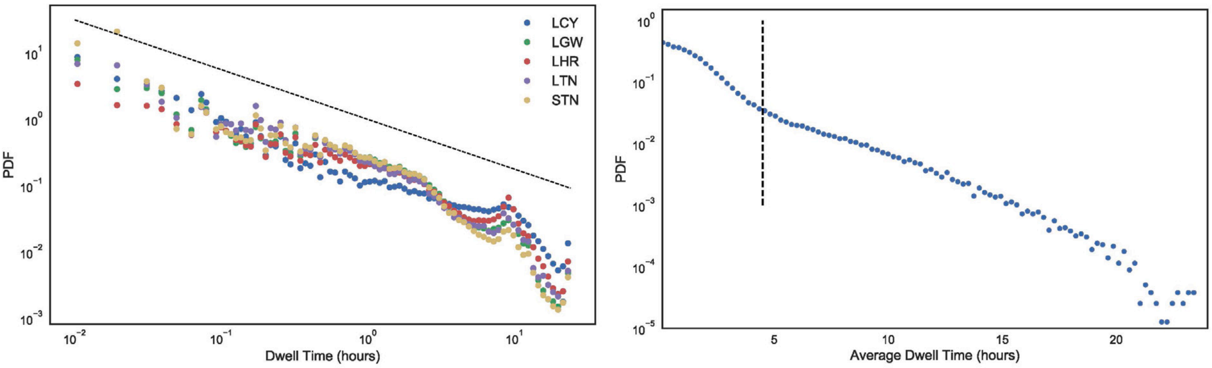
To characterize mobility from and to the airports, we aggregate a large dataset of anonymized GPS records in the UK generated by opted-in users of mobile applications in order to estimate distance and ground access times to/from airport facilities. The data was provided by Cuebiq Inc. The minimal spatial units considered in the study are middle layer super output areas (MSOAs), which represent geographical divisions employed in the UK census and surveys for statistical purposes. Within this database, we analyze six months of mobility data in the UK, covering the period February to July 2017. The information used for this study includes the UTC timestamp, anonymized user ID, and latitude and longitude coordinates. First, we isolate all the trajectories of the anonymized users for which at least a single GPS data-point falls within the geographical boundaries of the six airports in the Great London area, as shown in Fig. 1(a) (IATA codes: LHR, LTN, STN, LGW, LCY, SEN). For this purpose, the airport areas are approximated by a polygon manually constructed over satellite maps. We further require at least 12 data points to fall within the polygon in order to ensure the anonymized user was not simply traveling near the airport area. This implicitly imposes a minimal duration of the order of one hour. This minimal duration may vary depending on the mobile phone operating system, but under ideal conditions, we have a data point approximately every 5 min for Android phone users.

Subsequently, we aggregate the behavior of the anonymized users at the census level. We associate each observed trip to the airport undertaken by an anonymized user with a regularly visited location (home or workplace) based on the MSOA, covering the whole of England and Wales. These MSOAs have a population of about 7200, on average, as can be seen in Fig. 1(b), ensuring further anonymization of the aggregated data. The MSOA associated to 'home' represents the largest number of data points recorded at nighttime (after 6 pm and before 8 am), whilst 'work' draws from the largest number of points recorded during working hours, between 9 am and 5 pm. Distances from the airports are then measured as the minimum distance in a local projection (British National Grid) between the MSOA polygons and a point within the airport polygon.

In order to correctly frame the passenger analysis using the London airports alone, we isolate and exclude the trajectory produced by the workers at the airport and aviation companies. Consequently, we introduce two different filtering conditions: (I) An anonymous user is considered a worker if he/she is seen at the airport at least once on each of three consecutive days; (II) An anonymous user is considered a worker if his/her average dwell time at the airport exceeds 4.5 h (with the further requirement


(a) London airports (b) England and Wales (7,201 MSOAs) and the six airports surrounding London

**Fig. 1.** The case study of the six London airports and England and Wales MSOAs.




(a) Distribution of the dwell times at five airports (Southend excluded) (b) Average dwell time of users in all airports. The dashed line indicates the threshold value, set at 4.5 hours

**Fig. 2.** Distribution of the individual dwell times.

of appearing on at least three different days). The rationale behind condition I is that it is unlikely that an air passenger flies three days in a row from/to the same airport. As for condition II, as the distribution of the individual dwell times is similar across the six airports, we evaluate the value of the threshold for condition II as the point where a visible change is observed in the statistical behavior across the distribution of the average dwell times for the whole user-base, as shown in Fig. 2. This selection identifies a fraction of about 5% of travelers as workers at the different airports (with the exception of LCY, where they are 11%). The workers represent 30%–50% of the total number of daily trajectories over the six months analyzed. As the workers stay in the airport for longer times, they ultimately produce the majority of data-points recorded in the MSOA. We also exclude from our analysis trips to airports that lie in the same MSOA associated with the user's home or workplace because it falls below the resolution we set to ensure full anonymization of our results.

Once workers are identified and excluded from our analysis, we define two categories of users. The first category is *local residents*, who are observed across the United Kingdom for at least 40 days in the 6 months of our analysis and are observed inside the MSOA associated to home and workplace for at least 10 days. The second category is *potential tourists*, who are observed for less than 15 days. The identification of an anonymized user as a local UK resident appears to be reliable due to the abundance of data. Conversely, those labeled potential tourists are most likely a mix of tourists and local users who turn on their mobile applications sparsely. For this reason, to classify a user as a tourist in the London area, we further require the anonymized users to be observed in the data for at least 3 days and for at least 75% of the range of days between the first and last observation. Moreover, these users should be

**Table 1**
Descriptive statistics of airports in case study.

| Airport | Distance to London city center (kilometers) | Trip time by car (minutes) | Trip time by fastest public transport mode (minutes) | Passengers[a] (millions) | Air transport movements[a] (thousands) |
|---|---|---|---|---|---|
| LCY | 16 | 40 | 45 | 4.53 | 76.5 |
| LGW | 70 | 80 | 33 | 45.55 | 283.4 |
| LHR | 27 | 40 | 40 | 77.99 | 476.2 |
| LTN | 56 | 64 | 70 | 15.99 | 106.1 |
| SEN | 69 | 82 | 120 | 1.09 | 11.3 |
| STN | 66 | 72 | 100 | 25.90 | 172.6 |

[a]Source: CAA (2017).

**Table 2**
Descriptive statistics of the MSOAs.

| Statistic | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|
| Population | 7787.2 | 1599.6 | 2203 | 16,342 |
| Cars per capita | 0.5 | 0.1 | 0.1 | 0.8 |
| Bus stops count | 36.3 | 25.6 | 0 | 252 |
| With coach station | 17% | | | |
| With metro station | 6% | | | |
| With rail station | 27% | | | |
| Net weekly income (£) | 602.5 | 137.4 | 280 | 1400 |
| Area (km$^2$) | 21.0 | 53.1 | 0.3 | 1117.2 |
| Population density (per km$^2$) | 3218.0 | 3431.3 | 5.7 | 24,720.9 |

observed in proximity to at least two of the four most visited touristic locations in London (Tower of London, Buckingham Palace, Madame Tussauds, and the London Eye). Foreign travelers who do not meet the potential tourists' criteria are excluded from the data. Those could include, for example, business travelers. Finally, for every day in which an anonymized user was viewed at an airport, we estimate the time interval between the last point within the home and workplace MSOA (if recorded) and the arrival at the airport, or vice versa in the case of observed trips from the airport. We then associate the trip with the MSOA (home or work) that was temporally closer to the airport. For each MSOA, we record the number of trips and the distribution of the recorded travel times. Since there is no guarantee that the time between the last observation around the home (or work) and the observation at the airport perfectly correspond to the traveler's journey from/to the airport, we do not use the average travel times but rather the shape of the distribution recorded in the form of quartiles. These flows and travel times are subsequently connected to a set of socio-economic data extracted from the British Census at the MSOA level, in addition to statistics on the local accessibility of the UK public transportation network collected from openly available data (Gallotti and Barthelemy, 2015), in order to feed the regression analysis.

After creating the dataset, we perform data validation by comparing the processed information to the demand data published by the Civil Aviation Authority (CAA) and to the travel times recorded in the Google Maps API tool, as described in Section 2.3.

### 2.2. Case study of Greater London

The case study consists of six airports in the Greater London area: Heathrow (LHR), Gatwick (LGW), London City (LCY), Luton (LTN), Stansted (STN), and Southend (SEN), all of which currently serve scheduled airline services. Table 1 presents distance and trip times to central London, as well as the demand served in 2017. We make use of the geographical split of England and Wales into 7201 MSOAs; relevant socio-economic summary data are described in Table 2 and their correlations are presented in Table 3. From the raw data, it is clear that there is substantial variation in the dataset, with population sizes ranging from 2200 up to 16,300 inhabitants, and from MSOAs of one-third of a square km in London up to 1120 km in rural areas. Net weekly incomes also vary considerably, with average incomes in some MSOAs equal to half the mean weekly salary (£600), compared to more than double the average salary in wealthier areas. To explore the effects on ground accessibility, we also focus on vehicle ownership and its correlation with other characteristics. As expected, the regions with an underground or metro service tend to have lower levels of vehicle ownership. Higher population densities are also highly negatively correlated with ownership levels. Furthermore, bus stops are located more frequently in zones without a metro service, which could explain the positive correlation with car ownership. For this reason, we regress demand with respect to public transportation options and the lack thereof in order to explore the sensitivity of the other variables.

We analyze the demand of both locals and tourists for each of the six airports drawing from all MSOAs in England and Wales. In total, the dataset includes 104,158 trips by locals and 4580 trips by tourists in the six month period from January to July 2017. The aggregated data is presented in Table 4. We note that although the dataset represents less than 0.1% of the total trips, the sample is still substantially larger in comparison to any classical survey undertaken for such purposes. London Heathrow attracts the largest market share and serves more intercontinental routes than the other airports, which are more likely to serve domestic and European routes. As can be seen in Table 4, tourists are split between the more business-oriented, who opt for London Heathrow, and the

**Table 3**
Correlations between MSOA variables.

| | Population | Cars per capita | Bus stops count | Coach station | Metro station | Rail station | Income | Density |
|---|---|---|---|---|---|---|---|---|
| Population | 1 | | | | | | | |
| Cars per capita | −0.21 | 1 | | | | | | |
| Bus stops count | 0.12 | 0.42 | 1 | | | | | |
| Coach station | 0.09 | 0.02 | 0.20 | 1 | | | | |
| Metro station | 0.08 | −0.24 | −0.14 | 0.03 | 1 | | | |
| Rail station | 0.09 | 0.04 | 0.02 | 0.17 | 0.04 | 1 | | |
| Income | 0.01 | 0.35 | −0.12 | −0.08 | 0.10 | 0.07 | 1 | |
| Density | 0.18 | −0.73 | −0.50 | −0.11 | 0.23 | −0.09 | 0.03 | 1 |

**Table 4**
Demand distribution and market share per airport according to analysis.

| Airport | Locals trip demand | Locals market share | Tourists trip demand | Tourists market share | Total demand | Total market share |
|---|---|---|---|---|---|---|
| LCY | 9 203 | 0.09 | 91 | 0.02 | 9 294 | 0.09 |
| LGW | 27,379 | 0.26 | 1143 | 0.25 | 28,522 | 0.26 |
| LHR | 39,155 | 0.38 | 2143 | 0.47 | 41,298 | 0.38 |
| LTN | 10,899 | 0.10 | 175 | 0.04 | 11,074 | 0.10 |
| SEN | 793 | 0.01 | 10 | 0.00 | 803 | 0.01 |
| STN | 16,729 | 0.16 | 1020 | 0.22 | 17,749 | 0.16 |
| Total | 104,158 | 1.00 | 4582 | 1.00 | 108,740 | 1.00 |

leisure travelers, who are more likely to use Stansted compared to the locals. Both groups are equally likely to use Gatwick, but the remaining airports, namely, London City, Luton and Southend, tend to serve the local population.

### 2.3. Data validation

Validating the accuracy of the dataset is an important element of the analysis (Khan et al., 2014). In Fig. 3, we compare the dataset to the UK Civil Aviation Authorities (CAA) dataset from 2017 (CAA, 2017), after removing connecting passengers. The percentage of connecting passengers per airport is reported in the CAA Passenger Survey Report for 2016 (CAA, 2016).[2] Fig. 3 clearly shows that the annual market share across airports is very similar to that estimated based on the ICT data. However, there is a slight bias towards coverage of the more European business-oriented passengers than the general public on average. Consequently, Gatwick is slightly under-represented whilst LCY is over represented in the ICT dataset. In addition, we compare the median measured access times from each MSOA in the data analyzed to that of the Google Maps API data. In Fig. 4(a), we present the Google Maps estimation of travel time for the rush-hour compared to the estimated travel times for the 15,981 MSOA-airport pairs that were collected. It can be seen that for well covered MSOA-airport pairs, namely, those with at least 20 observations, the median measured times and Google times lie approximately on the 45° line. In Fig. 4(b), we compare the two datasets at an off-peak time (4 am) using the 25th percentile estimated access time, which seems reasonably well-matched. Finally, in Fig. 4(c), we compare the 75th percentile from the dataset to that of Google under the assumption that this represents public transport trips. Fig. 4(c) shows greater variability compared to Figs. 4(a) and 4(b), but still with significant estimators.

### 2.4. Consumer behavior analysis

In order to explore the characteristics that affect air travel demand, we first apply an ordinary least squares (OLS) regression. We aggregate demand according to origin–airport pairs, summed over both directions. Almost two thirds of origin–airport demand data pairs contain zero demand. In order to handle a large proportion of zero demand, we perform two regression analyses, similar to that of Fletcher et al. (2005). We first perform a binary logit regression for the complete dataset that measures the drivers effect on the existence of demand. These drivers include socio-economic characteristic of the MSOA (e.g., population size, average income level) and airport market power over the MSOA (e.g., its size, represented by the number of annual air traffic movements). To capture the role of the catchment area and potential overlapping areas, we include the travel time to the nearest alternative airport (which is the second closest for the MSOA-airport observation covering the closest airport). We also include a dummy which indicates whether the MSOA is served by additional airports within a 60 min radius. The last group of drivers are the travel options such as travel time to the airport, car ownership levels and public transport options. The logit regression equation is shown in Eq. (1).

$$\mathrm{P(flying)}_{i,j} = \frac{1}{1 + \exp\left(\gamma_0 + \sum_{k=1}^{K} \gamma_k X_{k,i,j}\right)}, \tag{1}$$

[2] As this report does not include Southend airport (SEN), we assume that no connecting passengers were served.

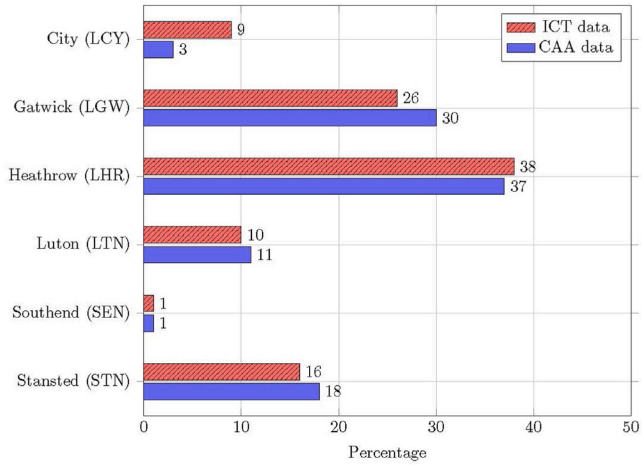
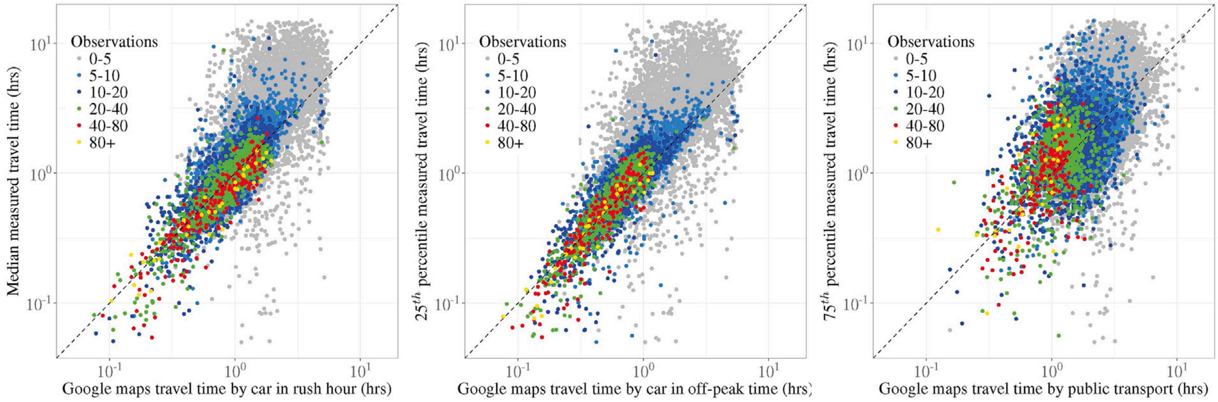



**Fig. 3.** Annual passengers and market share for 2017 compared to case study analysis.




(a) Access by car during rush hour ($R^2 = 0.57$) (b) Access by car during off-peak hours ($R^2 = 0.65$) (c) Access by public transportation ($R^2 = 0.27$)

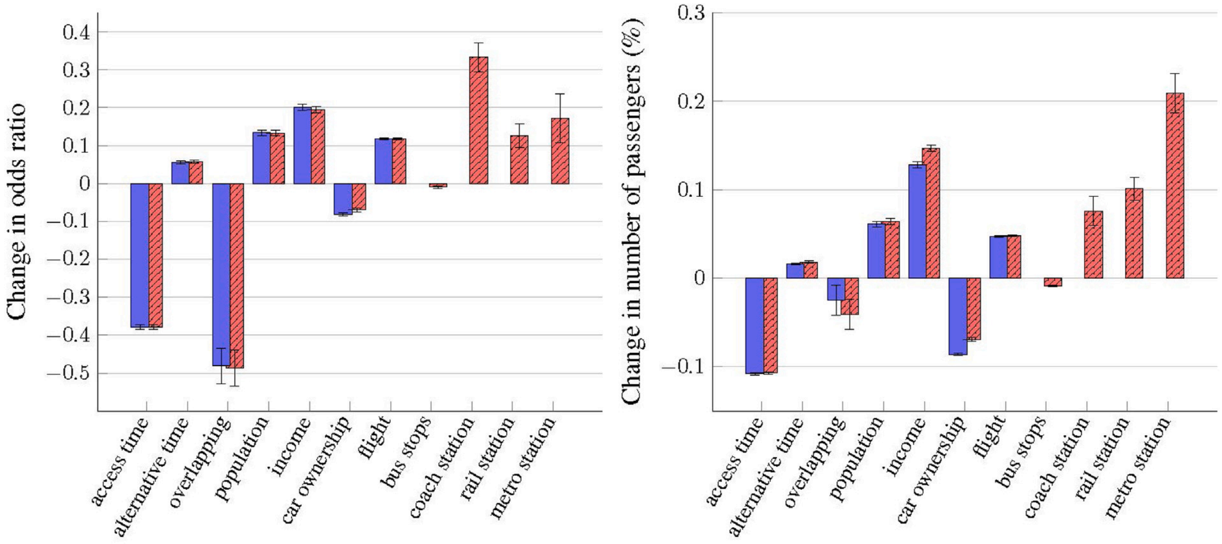
**Fig. 4.** Validation of the observed access times.

where P(flying)$_{i,j}$ is the probability of demand for airport $j$ from MSOA $i$, $X_{k,i,j}$ is the $k$th characteristic out of $K$ of MSOA $i$ with respect to airport $j$ (e.g., distance), $\gamma_k$ is the coefficient of characteristic $k$, and $\gamma_0$ is the intercept. In a second regression, we analyze only those origin–airport pairs with positive demand, as shown in Eq. (2).

$$\ln(\text{demand})_{i,j} = \eta_0 + \sum_{k=1}^{K} \eta_k \ln X_{k,i,j}, \tag{2}$$

such that the $\eta$ coefficients represent the elasticity of demand with respect to the relevant characteristic. Both regressions were performed twice, with and without public transportation capacities, for the purpose of sensitivity analysis.

We analyze the origin–airport characteristics that affect travel demand in general. The aggregated data cover 43,206 origin–airport demand pairs, of which around 65% contain zero demand. The binary logit regression presented in Eq. (1) for the entire dataset measures the variables' impact on the existence of demand. In addition, we perform an OLS regression of pairs with non-zero demand, presented in Eq. (2), that analyzes the characteristics affecting the magnitude of demand. The results of the two regressions are presented graphically in Fig. 5 and numerically in Table C.10 in Appendix C. We note that in both regressions, the non-binary independent variables are log transformed multiples of 10; therefore, in the logit regression (denoted models (a1) and (a2)), we estimate the change in the log odds of demand per 10 percent change in the continuous independent variables or binary variable. In the OLS regressions (denoted models (b1) and (b2)), we obtain the percentage change in demand for a 10 percent change in the continuous independent or binary variables. The results of models (a1) and (a2) (blue and red columns in Fig. 5(a), respectively) suggest that a 10% population growth in a MSOA (e.g., due to an increase in density) increases the demand odds by $e^{0.134} - 1 \approx 14\%$. The travel time to the nearest airport is, as expected, positively associated with the probability of flying from a particular airport from a specific MSOA. Increasing the relative travel time between an MSOA and an airport by 10%, decreases the demand odds




(a) logistic regressions (pseudo $R^2 = 0.390, 0.391$) (b) Ordinary least square regression (adjusted $R^2 = 0.542, 0.549$)

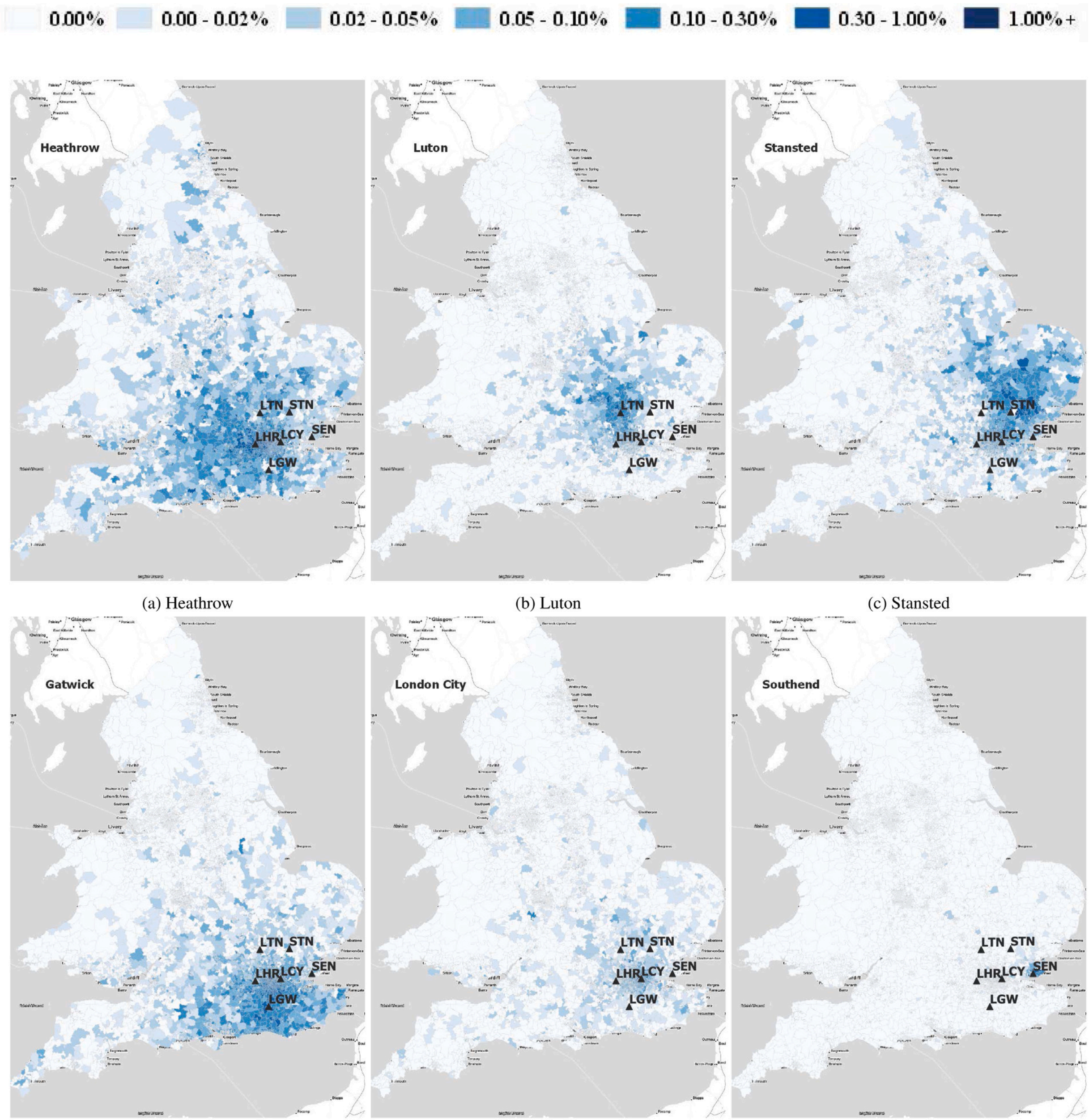
**Fig. 5.** Regression results with models (a1) and (b1) in blue and models (a2) and (b2) in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

by $1 - e^{-0.297} \approx 31\%$ percent. This highlights the importance of access times. Furthermore, in terms of access modes, rail, metro and intercity coach services increase the probability of positive demand between 10 and 40%. However, bus services are negatively associated with the probability of demand. We assume that areas located within Greater London are better characterized by rail and metro than the rest of England. For the residents in a MSOA served by multiple airports, the overlapping catchment area has a negative impact on demand for each of the respective airports because residents enjoy more options, indicating some competition between the airports, at least at the service level, for example by offering similar destinations.

We estimate the drivers of demand from MSOA to airports for which demand is positive and the results of models (b1) and (b2) are presented in Fig. 5(b) (blue and red columns, respectively). As expected, because aviation is considered a discretionary-derived demand, the higher the average MSOA income, the higher the expected trip demand levels. With respect to the catchment areas, the coefficient of the 'overlapping' variable, which equals one if more than one airport is deemed to be within the catchment area of a MSOA, is negative suggesting competition reduces demand. The greater the distance to the next nearest alternative facility, the higher the demand for the specific airport, as demonstrated by the 'time to 2nd airport' variable. Again, access time plays an essential role in determining the demand. With respect to car ownership, the regression results suggest that 10% higher ownership leads to 7 to 9% lower demand. This may be due to (i) car owners choosing to drive for short-distance trips instead of taking a flight, or alternatively, (ii) due to MSOAs with higher population densities leading to lower car ownership shares, particularly in London. In contrast, improved access to rail and intercity bus (coach) services increases demand for trips by 8 to 10% and the underground/metro stations show the strongest effect, with an approximately 22% increase in demand. Moreover, the inclusion of public transportation alternatives in the regression did not change the results, confirming their robustness with respect to their effect on aviation demand.

### 2.5. Competition across catchment areas

Assuming that the ICT dataset is sufficiently representative, we create catchment area maps from the GPS data compiled for the local inhabitants only. Figs. 6(b) through 6(g) present the catchment areas for each of the six airports included in the case study. The visualization of the catchment areas around the airports clearly shows substantial differences, which are partly due to the dissimilar services offered in terms of final destinations, information that is missing from the current ICT dataset. For example, LHR and to a lesser extent LGW act as intercontinental gateways for the British Airways network, a member of the One World alliance. Consequently, the catchment area of LHR covers many of the MSOAs across England and Wales, whereas STN and LTN serve mostly low-cost carriers, namely Ryanair and Easyjet, thus offering services of a dissimilar type to that of LHR. Luton's catchment area is mostly to the north of the airport location, whereas Stansted serves a narrow, long corridor to the north and east. However, there is an overlapping area to the east of Luton and west of Stansted from which local residents are served by both airports. Gatwick airport serves mostly the south of England, whereas Southend and London City airports serve the immediate area around their respective locations. It has been argued in the literature that airports in multi-airport regions tend to differentiate themselves in an attempt to avoid head-on competition (Fröhlich and Niemeier, 2011). Based on the catchment area maps of Fig. 7 which highlights the level of market dominance of an airport over a zone it seems that airports tend to be monopolistic markets, except for Luton




(a) Heathrow (b) Luton (c) Stansted

(d) Gatwick (e) London City (f) Southend

**Fig. 6.** Airport catchment areas for local inhabitants.

and Stansted that have a relatively small overlapping catchment area for local residents. It is further emphasizing the importance of ground transport access on airport competition.

In Fig. 8, we show the difference in the ground access travel time between each pair of the six airports analyzed (e.g., LHR-LCY in the top left panel). The time difference measures the gap in trip time to each airport for inhabitants residing in one MSOA such that positive values represent a shorter distance to the second airport. A cloud of points represents a single MSOA area, and each solid line a running average (red for the first airport, blue for the second airport). The Southend (SEN) airport was excluded from this plot due to the limited amount of data gathered. A comparison of all the curves highlights that many couples display almost symmetrical running averages, indicating a substantial overlap in the potential markets. This is particularly true for the cases of the LHR-LGW and LTN-STN pairs.

In summation, the theoretical assumption that a catchment area around an airport is a circle of approximately 100–150 km would appear to be an inaccurate depiction of reality (Adler and Liebert, 2014; Sun et al., 2021), at least in the case of the London region.

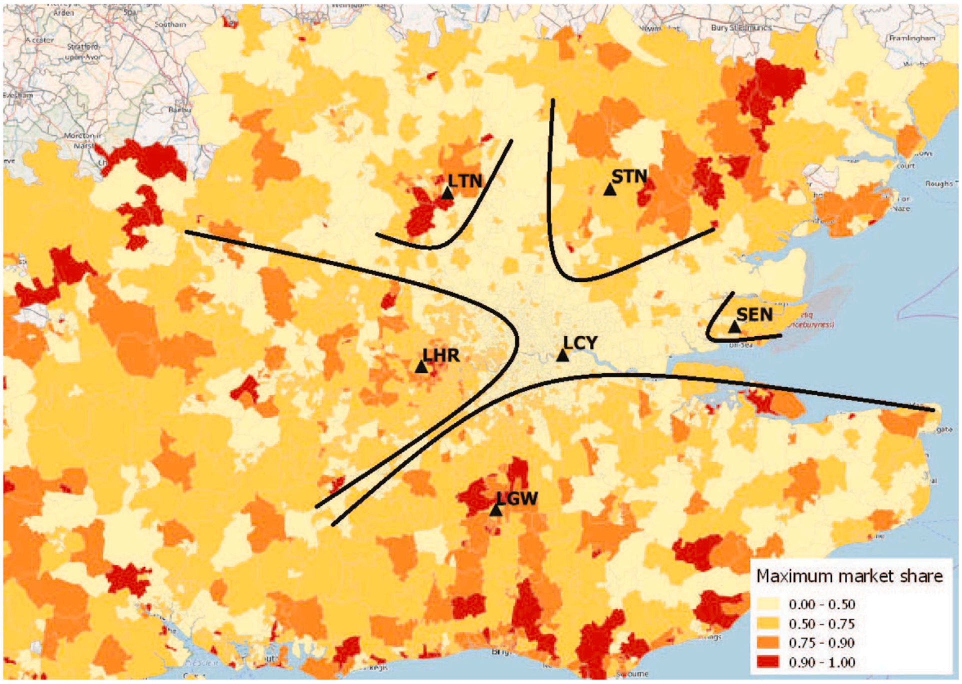



**Fig. 7.** Intensity map.

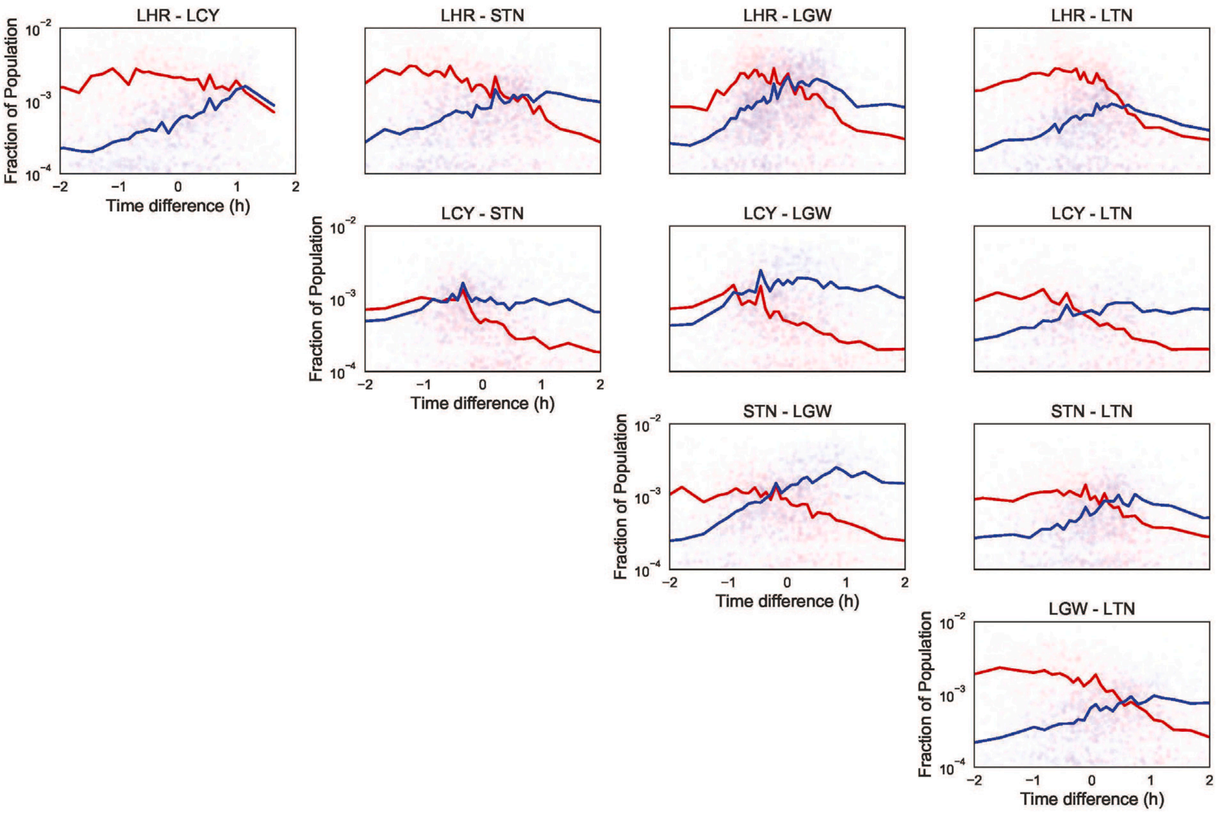



**Fig. 8.** Empirical market ground access trip times between every airport pair. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 5**
% of overlapping destinations between airports.

| Airport | Number of destinations | LHR | LGW | STN | LTN | LCY | SEN |
|---|---|---|---|---|---|---|---|
| LHR | 495 | 1.00 | | | | | |
| LGW | 248 | 0.77 | 1.00 | | | | |
| STN | 165 | 0.56 | 0.56 | 1.00 | | | |
| LTN | 91 | 0.65 | 0.69 | 0.68 | 1.00 | | |
| LCY | 45 | 0.73 | 0.67 | 0.58 | 0.42 | 1.00 | |
| SEN | 20 | 0.65 | 0.80 | 0.75 | 0.65 | 0.45 | 1.00 |

The distinct catchment areas may suggest that airport-airline pairs compete by offering common destinations. Some confirmation comes from the data about shared final destinations between each pair of airports as depicted in Table 5. However, it does not provide an understanding of other potential collusive issues, such as price coordination. For that, we present in the following section a theoretical analysis that explains levels of competition. We focus our case study on the LTN-STN pair due to the 68% overlap in final destinations, as shown in Table 5), and the similar airline business models, which would suggest the potential for gains from collusion.

## 3. Analytical model and producer behavior

In this section, we develop a catchment area game drawing insight from Section 2. In Section 3.1, we describe the behavior of the producers serving the market and discuss the potential impact of vertical and horizontal collusion. In Section 3.2 we solve the game analytically for a simple symmetric game, which we solve and extend to a numerical, non-symmetric sensitivity analysis in Section 3.3. Finally, we analyze the Stansted–Luton case in Section 3.4.

### 3.1. Catchment area game

We develop a catchment area game featuring two airports and two airport–airline pairs. We extend and adapt the Barbot (2009) model of vertical collusion between airport–airline pairs in order to explore horizontal collusion between airports, as depicted in the top graph of Fig. 9. The bottom graph part is an enlarged version of the empirical ground access analysis of LTN-STN shown in Fig. 8. The fraction of the total population traveling from/to one of the two airports is a function of the difference in time required to reach the airport from the origin MSOA to Luton (LTN) and Stansted (STN). Points represent single MSOAs. The dashed line is the running average, and the colored margins are the standard error of the mean. The crossover point represents the indifference point $I$, defined in the top graph.

To construct the catchment area game, first consider the top graph of Fig. 9, where airports $L$ and $S$ are located within $d$ minutes drive from one another. Airline $E$ operates at airport $L$ and airline $R$ operates at airport $S$. Each airport–airline pair has a monopolistic catchment area, hence the superscript $m$. For the pair $(L, E)$, this area is located to the west of the airport, and for the pair $(S, R)$, to the east. A passenger located within $x^m_L$ minutes from airport $L$ will choose to fly from this airport if the overall cost of service, i.e., the sum of the access costs and airfare, is lower than $V_E$, which represents the passenger's willingness to pay for the service. The potential demand for the $(L, E)$ service in their monopoly region is a function of the number of individuals living within $x^m_L$ minutes from airport $L$. The area between the two airports is the "competitive" area, hence the superscript $c$. We assume that distances $x^c_L$ and $x^c_S$ minutes from the airports are the same distance from the airport as $x^m_L$ and $x^m_S$, respectively. However, $x^c_L$ and $x^c_S$ may overlap, creating an overlapping catchment area between the airports. Denoting $0 \le \chi \le 1$, a passenger located at point $I$ within the overlapping catchment area is $d\chi$ minutes from airport $L$ and $d(1-\chi)$ minutes from airport $S$ and is indifferent (hence the $I$) between the two airports. Based on this simplification, the deterministic choice model assumes that a passenger located to the west of $I$ chooses airport $L$ and a passenger located to the east of $I$ chooses airport $S$. The access time $x^m_L$ is estimated according to Eq. (3).

$$V_E = tx^m_L + p_E, \quad \text{which yields:} \quad x^m_L = \frac{V_E - p_E}{t}$$
$$\text{and for } S\text{:} \quad x^m_S = \frac{V_R - p_R}{t}, \tag{3}$$

where $t$ is the access cost per minute, $p_E$ and $p_R$ are the airfares of airlines $E$ and $R$, respectively. A summary of the parameters and decision variables is presented in Table 6.

For the competitive catchment area, the equilibrium ratio is defined as in Eq. (4).

$$V_E - (d\chi t + p_E) = V_R - (d(1-\chi)t + p_R) \quad \text{which yields:}$$
$$d\chi = \frac{V_E - V_R + td + p_R - p_E}{2t}, \tag{4}$$

where $\chi$ is the fraction of $d$ that is captured by airport $L$ ($(1-\chi)$ is captured by airport $S$).

Assuming that the potential passengers are spread evenly across the catchment area, then airport $L$'s demand to the west of the airport, $y^m_L$, is defined as in Eq. (5).

$$y^m_L = r^m_L x^m_L \quad \text{and similarly} \quad y^m_S = r^m_S x^m_S, \tag{5}$$

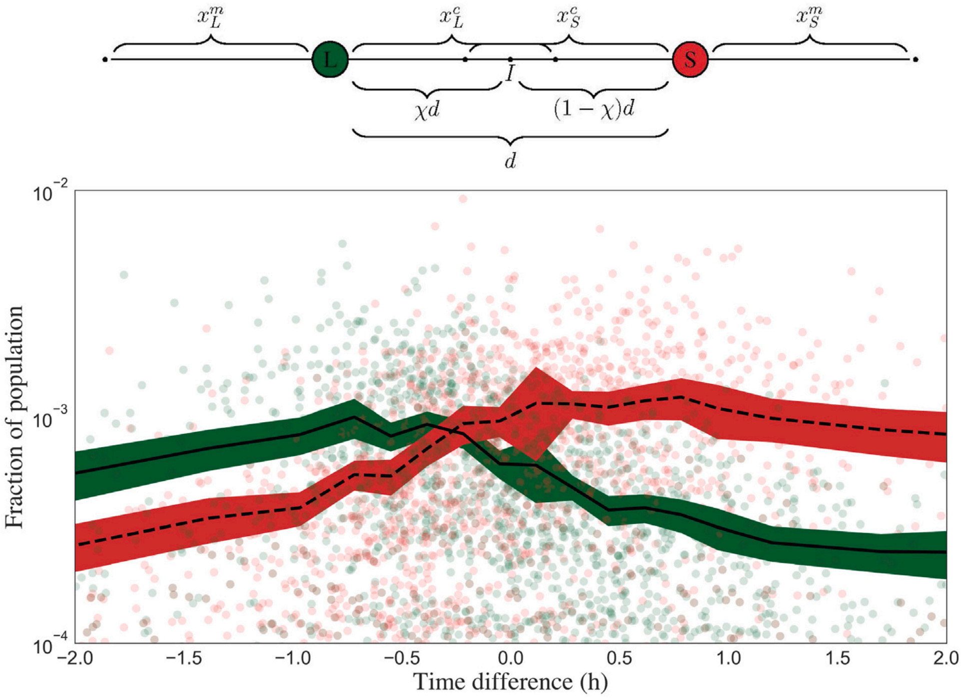



**Fig. 9.** The catchment area game of the empirical estimation of the market competition between Luton and Stansted, produced by the average ground access trip times. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 6**
Notation.

| Type | Symbol | Meaning |
|---|---|---|
| Decision variables | $p_i$ | Airfare of airline $i = E, R$ |
| | $p_j$ | Airport charge at $j = L, S$ |
| | $k_j$ | Commercial prices at airport $j$ |
| Parameters | $r^c$ | Potential passengers per minute drive from each MSOA in the competitive area |
| | $r_j^m$ | Potential passengers per minute drive from each MSOA in $j$'s monopoly area |
| | $d$ | Travel distance (in minutes) between airports |
| | $t$ | Ground access cost |
| | $V_i$ | Willingness to pay to travel with airline $i$ |
| | $\beta_j$ | Price sensitivity of airport amenities $j$ |
| | $\alpha_j$, $\delta_j$ | Parameters of airport $j$'s cost function |
| | $c_i$ | Airline $i$'s cost per passenger |
| Outcomes | $I$ | Indifference point between airports |
| | $\chi$ | Share of passengers choosing one airport over the other |
| | $x_j^m$ | Airport $j$ catchment area size in the monopoly area |
| | $x_j^c$ | Airport $j$ catchment area size in the competitive area |
| | $y_j^m$ | Airport $j$ demand in the monopoly area |
| | $y_j^c$ | Airport $j$ demand in the competitive area |
| | $y_j^k$ | Airport $j$ demand for commercial amenities |
| | $tc_j$ | Airport $j$ total cost |
| | $\pi_i$ | Airline $i$ profit |
| | $\pi_j$ | Airport $j$ profit |
| | $\pi_{i+j}$ | Airline $i$ and airport $j$ aggregate profit |

where $r_L^m$ and $r_S^m$ are the demand per time unit distance west of airport $L$ and east of airport $S$, respectively. Consequently, the demand for airports $L$ and $S$ from the competitive area are defined as:

$$y_L^c = \min\left(r^c d\chi, r^c x_L^c\right) \quad \text{and} \quad y_S^c = \min\left(r^c d(1-\chi), r^c x_S^c\right), \tag{6}$$

where $r^c$ is the demand per unit distance in the competitive area. Based on the ICT data presented in Section 2, we know that the overlapping area exists, hence in the analysis we explicitly assume market covering for the competitive area, i.e., Eq. (6) can be rewritten as $y_L^c = r^c d\chi$ and $y_S^c = r^c d(1-\chi)$ and that there is a non-negative demand for both airports in their monopoly area and in the competitive area.

The passengers also consume amenities at the airport. The demand for amenities is a function of the number of passengers, the prices and price sensitivity, as shown in Eq. (7).

$$y_L^k = (y_L^m + y_L^c) - (\beta_L k_L) \quad \text{and} \quad y_S^k = (y_S^m + y_S^c) - (\beta_S k_S), \tag{7}$$

where $k_L$ and $k_S$ are the average prices for amenities in airports $L$ and $S$, respectively, and $\beta_L$ and $\beta_S$ represent the price sensitivity. Note that according to Eq. (7), the result is that the optimal average price for an amenity is such that half the passengers will purchase a product or service at the airport.[3]

For the cost structure at the airport, we assume a linear cost function. Hence, we formulate the total cost at the airport as follows:

$$tc_L = \alpha_L(y_L^m + y_L^c) + \delta_L \quad \text{and} \quad tc_S = \alpha_S(y_S^m + y_S^c) + \delta_S, \tag{8}$$

where $\alpha$ and $\delta$ are parameters of the cost function. The airport profit is given by:

$$\pi_L = p_L(y_L^m + y_L^c) + k_L y_L^k - tc_L \quad \text{and} \quad \pi_S = p_S(y_S^m + y_S^c) + k_S y_S^k - tc_S, \tag{9}$$

where $p_L$ and $p_S$ are prices per passenger that the airports charge the airlines. The airline profit function is given by:

$$\pi_E = (p_E - c_E - p_L)(y_L^m + y_L^c) \quad \text{and} \quad \pi_R = (p_R - c_R - p_S)(y_S^m + y_S^c), \tag{10}$$

where $c_E$ and $c_R$ are each airline's operating cost per passenger.

Consequently, we define a two-stage, oligopolistic market where in the first stage, two airports simultaneously set charges to be paid by airlines. In the second stage, the airlines simultaneously set their airfares whilst accounting for the airport charges from the first stage and the cost of travel to the airport as a function of the furthest passenger's location within the catchment area. In addition, the airports set commercial prices to be paid by passengers in the second stage. This is a complete, imperfect information game, as presented in Fig. 10, where, given the mobility and economic data collected, we estimate the payoffs for eight market scenarios: (i) full competition, (ii+iii) single airport–airline vertical collusion, (iv) both airport–airline pairs vertically collude, and (v–viii) airport-airport horizontal collusion combined with scenarios (i) to (iv).

We solve the game using analytical and numerical solution methods. When using analytics, in order to maintain tractability, we solve the game for the scenarios in which both airport–airline pairs vertically collude or not (i.e., scenarios (i), (iv), (v), and (viii)). We assume symmetry, i.e., that the airports and airlines are of the same size and cost structure. In the numerical analysis, we solve a symmetric case study for all scenarios and then compare the results to the real-world, asymmetric case of two London airports with clearly identifiable overlapping catchment areas according to the results of the analysis described in Section 2.1. Specifically, we analyze Luton ($L$) and Stansted ($S$), which function as hubs for Eastjet ($E$) and Ryanair ($R$), respectively. The market specific characteristics are that Stansted is the larger of the two, and Ryanair's cost structure is lower than that of Easyjet. In the collusive scenarios, profits depend on the airport's bargaining power versus that of the hubbing airline. We solve the joint profit functions and split any additional gain or loss evenly between the airport and the airline, according to the Nash bargaining solution (Nash, 1950). We estimate the additional profit or loss in comparison to scenario (i), the purely competitive case, for scenarios (ii)–(iv), and in comparison to scenario (v), the horizontal collusion case, for scenarios (vi)–(viii).

### 3.2. Competition or collusion between producers

We solve scenarios (i), (iv), (v), and (viii) analytically under the assumption that the airports and airlines are of the same size and cost structure (see Fig. 10 and Appendix A). We also assume that the parameters and decision variables are non-negative. In other words, there are no subsidies and parameter values must be such that no airline or airport is so uncompetitive that it is driven out of business. Due to symmetry, we refer to airline $i$ and airport $j$. We first present the analysis from the producers' perspectives and then the consumers (see Appendix B for proofs).

Solving for airfares across the four scenarios, leads to the following ranking[4]:

$$p_i^{(iv)} \le p_i^{(viii)} \le p_i^{(i)} \le p_i^{(v)}, \quad i \in E, R. \tag{11}$$

This interesting result suggests that the purely competitive solution outcome does not lead to the lowest airfares as may be expected. Rather, airport–airline vertical collusion lowers airfares beyond the competitive outcome (scenario (i)). This phenomenon might be in part due to the elimination of the issue of double marginalization (Gaudet and Long, 1996). An alternative explanation relates to the behavior of airport management who lower airport charges in the expectation that airlines reduce airfares which in turn attracts more passengers. Any deviation by an airline that increases airfare will result in lower demand and increased airport charges. On the other hand, as expected, airport-airport horizontal collusion leads to higher airfares. Perhaps surprisingly, the case of complete collusion (scenario viii) also leads to lower airfares compared to the purely competitive market (scenario i). In other words, the downward

[3] The profit from the terminal side is $\pi^k = ((y^m + y^c) - \beta k)\,k$. The optimal price leads to demand $k^* = (y^m + y^c)/2\beta$. $y^{k*} = (y^m + y^c)/2$.

[4] For proof see Appendix B.2.

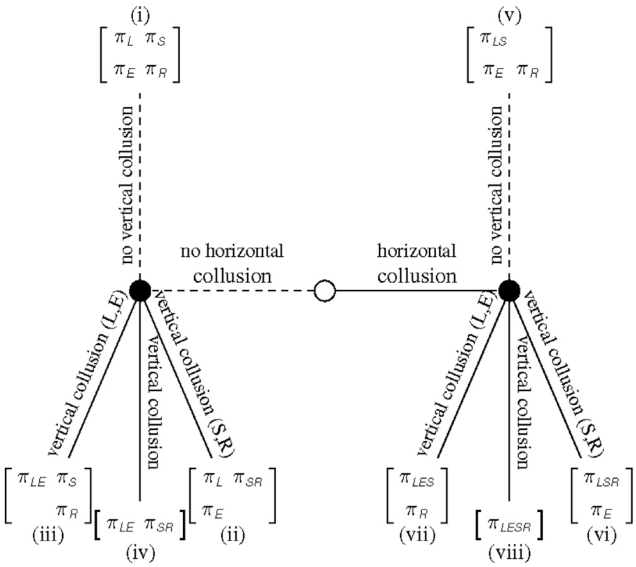



**Fig. 10.** The two-stage game.

pressure on airfares from vertical collusion outweighs the negative impact of horizontal collusion, leading to lower airfares when all parties collude, as compared to the competitive outcome. This outcome might be explained by the monopoly power an airport enjoys over a restricted geographical area, hence, the contribution of horizontal collusion is rather limited compared to that of vertical collusion. Another potential explanation might be the characteristics of the model in which we assume that a single airline serves each airport. The level of demand is in reverse order to the airfare ranking shown in Eq. (11). Therefore, the least favorable outcome from the customers' perspective is horizontal collusion without vertical collusion (scenario v), which leads to the highest airfares amongst the scenarios analyzed. Full collusion is the second favorable outcome for consumers after vertical collusion.

From the profit perspective, we show that horizontal collusion increases the profits of the airports at the expense of the airlines, as compared to pure competition (for proof see Appendix B.1): $\pi_j^{(v)} \geq \pi_j^{(i)}$ and $\pi_i^{(v)} \leq \pi_i^{(i)}$. We note that in addition, aggregate profits under vertical collusion (scenario (iv)) are higher than those of the competitive market (scenario (i)): $\pi_{i+j}^{(iv)} \geq \pi_i^{(i)} + \pi_j^{(i)}$, for a higher access cost threshold as shown in Proposition 3. Finally, comparing scenarios (iv) and (viii) shows that the aggregate profits for the vertically colluding airline $i$ and airport $j$ under complete collusion ($\pi_{i+j}^{(viii)}$) are higher as compared to purely vertical collusion ($\pi_{i+j}^{(iv)}$), hence scenario (viii) is the Nash equilibria outcome: $\pi_{i+j}^{(viii)} \geq \pi_{i+j}^{(iv)}$.

### 3.3. Numerical analysis of the catchment area game

The numerical analysis of all the scenarios represents the case of Luton-Easyjet $(L, E)$ and Stansted–Ryanair $(S, R)$. The parameters collected for the numerical analysis are presented in Table 7. Before turning to the case study, we run the numerical analysis for a symmetric catchment area game and assume that the airports and airlines have the same size and cost structure. For this purpose, we compute the average values of the parameters presented in Table 7. The symmetric catchment area game results appear in Table 8(a) where in each cell, the four numbers represent the profits of airports on the top row and those of airlines on the bottom row. The left column is for the $(L, E)$ pair and the right one for the $(S, R)$ pair. Each airport decides whether to compete with its counterpart (represented by the four cells of scenarios (i)–(iv) in the top left corner) or to collude (represented by the four cells of scenarios (v)–(viii) in the bottom right corner). Each hub airport–airline pair decides whether to vertically collude (scenarios (ii)–(iv) and (vi)–(viii)) or not (scenarios (i) and (v)).

The analysis reveals that complete collusion (scenario (viii)) is the equilibrium outcome, involving both horizontal and vertical collusion, as shown analytically in Section 3.2. To explore the robustness of that result, we introduce asymmetry between the players with respect to the cost function of the airline and the potential market size. We find that for this case, if one of the airlines has at least 17% higher costs per passenger than the other airline, it is sufficient for the equilibrium to shift from (viii) to (vi), suggesting that the airport–airline pair that produces services at a lower cost might be reluctant to collude with the other pair. The same occurs when one airport monopoly catchment area is 27% denser in population per minute drive. Different airport costs are also a factor, however, because an airport cost structure includes a high share of fixed costs, the difference in variable costs would need to be three times higher in order to shift the equilibrium. From the broader perspective, this quantitative assessment suggests that regulators ought to carefully assess horizontal collusion between airports. In contrast, vertical collusion between airports and airlines might not need to be discouraged because it is also favorable for customers. On the other hand, we also note that vertical collusion could potentially deter the entry of new airlines, thus creating opportunities for weaker competitive outcomes.

**Table 7**
Parameters applied to the numerical case.

| Symbol | Measure units | Meaning | Value | References |
|---|---|---|---|---|
| $r_L^m$, $r_S^m$, $r^c$ | people | Potential passengers per minute drive from each MSOA | 120,000, 240,000, 80,000 | Dataset |
| $d$ | minutes | Distance in minutes drive between airports | 60 | Google maps |
| $t$ | £ | Cost per minute drive | 0.45 | AAA |
| $V_E$, $V_R$ | £ | Willingness-to-pay for trip | 92, 72 | Merkert and Beck (2017) |
| $\beta_L$, $\beta_S$ | | Commercial price sensitivity | 0.30, 0.60 ($\times 10^6$) | Del Chiappa et al. (2016) and Andreyeva et al. (2010) |
| $\alpha, \delta$ | | Parameters of the airport cost function | Luton: 4, $70 \times 10^6$<br>Stansted: 2, $160 \times 10^6$ | Bottasso and Conti (2012)<br>Financial reports[a] |
| $c_E$, $c_R$ | £ | Airline average cost per passenger | 52, 29 | Financial reports[b] |

[a]Luton: https://www.london-luton.co.uk/CMSPages/GetFile.aspx?guid=62eca0aa-5002-47ac-a364-be135ded21a9 Stansted: https://www.magairports.com/media/1416/annual-report-year-ended-31st-march-2017-mahl.pdf.

[b]Easyjet: http://corporate.easyjet.com/~/media/Files/E/Easyjet/pdf/investors/results-centre/2017/2017-annualreport-and-accounts-v1.pdf Ryanair: https://investor.ryanair.com/wp-content/uploads/2017/07/Ryanair-FY2017-Annual-Report.pdf.

**Table 8**
The numerical solutions. The left column in each cell represents the $(L, E)$ pair, and the right column the $(S, R)$ pair. The top row represents airport profits, and the bottom row airline profits (in millions of £)

**(a) The numerical symmetric case**

| | | | Airport S (with airline R): Without horizontal collusion, Without vertical collusion | Airport S (with airline R): Without horizontal collusion, With vertical collusion | Airport S (with airline R): With horizontal collusion, Without vertical collusion | Airport S (with airline R): With horizontal collusion, With vertical collusion |
|---|---|---|---|---|---|---|
| Airport L (and airline E) | Without horizontal collusion | Without vertical collusion | (i)<br>27 27<br>79 79 | (ii)<br>−8 73<br>60 125 | | |
| | | With vertical collusion | (iii)<br>73 −8<br>125 60 | (iv)<br>38 38<br>90 90 | | |
| | With horizontal collusion | Without vertical collusion | | | (v)<br>29 29<br>64 64 | (vi)<br>55 55<br>38 91 |
| | | With vertical collusion | | | (vii)<br>55 55<br>91 38 | (viii)<br>**54 54**<br>**90 90** |

**(b) Luton–Stansted case**

| | | | Airport S (with airline R): Without horizontal collusion, Without vertical collusion | Airport S (with airline R): Without horizontal collusion, With vertical collusion | Airport S (with airline R): With horizontal collusion, Without vertical collusion | Airport S (with airline R): With horizontal collusion, With vertical collusion |
|---|---|---|---|---|---|---|
| Airport L (and airline E) | Without horizontal collusion | Without vertical collusion | (i)<br>15 11<br>49 97 | (ii)<br>−2 74<br>40 160 | | |
| | | With vertical collusion | (iii)<br>48 −7<br>83 87 | (iv)<br>**29 55**<br>**64 141** | | |
| | With horizontal collusion | Without vertical collusion | | | (v)<br>15 12<br>42 90 | (vi)<br>55 52<br>26 130 |
| | | With vertical collusion | | | (vii)<br>34 31<br>61 74 | (viii)<br>50 47<br>77 125 |

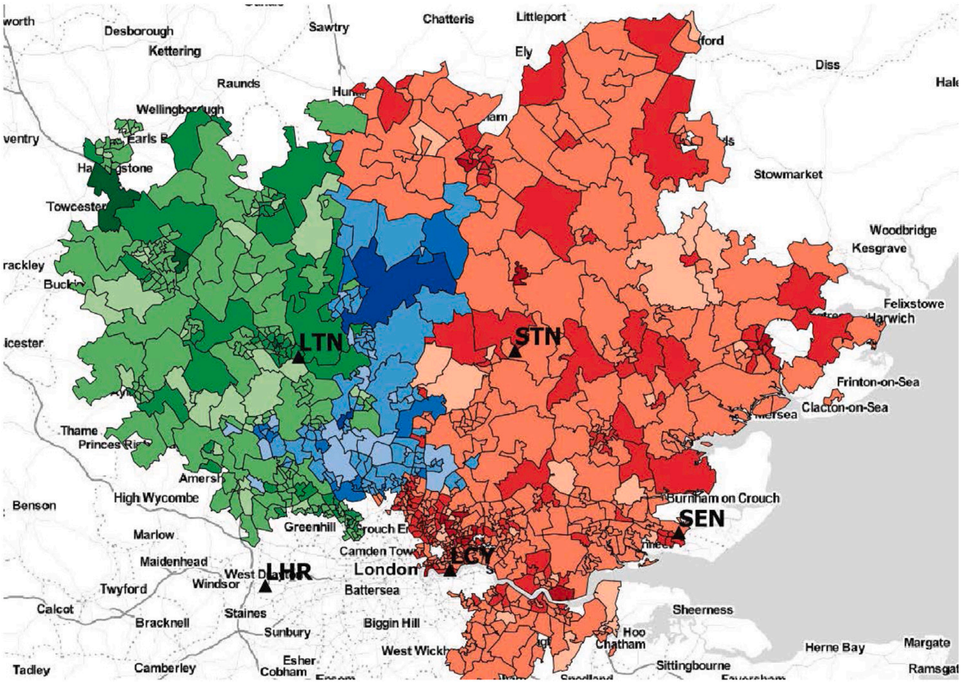



**Fig. 11.** Luton and Stansted catchment areas and overlapping region (blue zone). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We further note that the assumption of full market coverage in the overlapping area always leads to an interior solution, i.e. both airports serve strictly positive demand in the competitive catchment area. We also explored the option of a corner solution in which one airport–airline pair solely concentrates on extracting profits from the monopoly area, thus avoiding head-on competition for passengers. However, this consistently results in lower profits for the producers across all scenarios, suggesting that the internal solution is the unique equilibrium outcome regardless of the full coverage assumption. Indeed, this finding is mirrored in the results of the ICT data analysis which shows that the overlapping catchment area is fully covered and served by both airports (the blue area in Fig. 11).

### 3.4. Results of the Luton-Stansted catchment area game

We now turn to assess pricing competition between two London airports serving very similar, primarily leisure markets. Given the available road and public transport network services, Luton and Stansted have clearly defined separate catchment areas. Since most transit corridors are directed towards London, access times from the MSOAs east of Stansted to Luton and west of Luton to Stansted are long, despite the relative geographical proximity. The asymmetry between Luton-Easyjet and Stansted–Ryanair is both in the per-passenger cost of each airline and in the potential market size, where both differences are well above the thresholds found in Section 3.3. The data was collected from financial reports and shows that Ryanair cost per passenger are 60% that of Easyjet. By comparing the fiscal year of 2020, which was subject to many lockdowns due to the Covid19 pandemic and saw the plummeting in the numbers of passengers, to the year before, we see that Luton is characterized by smaller fixed costs with higher variable costs.[5] The potential demand is presented in Fig. 11 and consists of zones located up to an hour drive from Luton (green), Stansted (red), or both (blue); the darker the color of the zone, the greater the population size. The blue overlapping catchment area to the east of Luton and to the west of Stansted consists of 111 zones and approximately 800,000 residents, of which 54% chose Luton and 46% Stansted, according to the ICT data. The green region contains 246 MSOAs with a total of two million inhabitants, who require less than one hour to travel to Luton and more than one hour to reach Stansted. Approximately 80% of the passengers residing in this region chose to fly through Luton. In contrast, the red region consists of 479 MSOAs with around four million inhabitants, of which 85% of the passengers chose to fly from Stansted. Fig. 11 also confirms the assumption that the competitive area between the two airports is fully covered by the two airports i.e., everyone travels from these MSOAs to one of the two airports.

[5] From January to December 2020, around 5 million passengers went through Luton airport, and from April 2020 to March 2021, 3 million passengers went through Stansted. Comparing the fiscal years 2020 and 2019 allowed us to estimate the variable and fixed costs for each airport as depicted in Table 7.

**Table 9**
Real world data estimation compared to the catchment area game results.

| Variable | Units | Real world estimation | | Scenario (i) | | Scenario (iv) | | Scenario (v) | | Scenario (viii) | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | LTN | STN | LTN | STN | LTN | STN | LTN | STN | LTN | STN |
| Local passengers | millions | 10 | 16 | 4 | 8 | 9 | 17 | 4 | 8 | 7 | 16 |
| Airport profit | million £ | 27 | 57 | 15 | 11 | 29 | 55 | 15 | 12 | 50 | 47 |
| Airline profit | million £ | 53 | 162 | 49 | 97 | 63 | 141 | 42 | 90 | 77 | 125 |
| Average airfare | £ | 62.9 | 45.5 | 84.5 | 62.3 | 66.5 | 44.5 | 85.6 | 62.9 | 71.9 | 47.4 |
| Airport charge per passenger | £ | 6.2 | 6.2 | 20.8 | 20.8 | 7.4 | 7.3 | 22.8 | 21.9 | 11.7 | 9.4 |
| Commercial revenue per passenger | £ | 5.2 | 5.4 | 3.5 | 3.2 | 7.5 | 7.2 | 3.2 | 3.1 | 6.1 | 6.6 |

The results are depicted in Table 8(b). Comparing pure competition (i) and airport–airline vertical collusion (iv), the Nash equilibrium outcome suggests that both airport–airline combinations will be more profitable when colluding (scenario (iv)). In line with the analytical fare rankings, this is also the preferable outcome for passengers, because this equilibrium offers the lowest airfares hence, the highest utility compared to the other three scenarios. In the scenarios involving horizontal collusion between airports ((v)–(viii)), we see that both airport–airline pairs still prefer to vertically collude, which leads to the highest profits under the complete collusive outcome (scenario (viii)). Consequently, when comparing the results in both quadrants, we conclude that vertical collusion is likely. However, a comparison of the cases where both airport–airline pairs vertically collude ((iv) and (viii)) reveals that only the smaller $(L, E)$ pair strongly prefers complete collusion (viii), whereas the larger $(S, R)$ pair is better off without horizontal collusion (iv). Since collusion requires cooperation between all relevant parties, the equilibria outcome will likely be vertical, but not horizontal, collusion (scenario (iv)). Under the Nash bargaining approach, any additional profits or losses from a partnership are shared equally between all parties, in this case between the four producers. We note that the additional profits achieved in scenario (viii) compared to that of the disagreement point scenario (v) are not sufficient to encourage the larger pair to participate. If we were to assume that $(S, R)$ have stronger bargaining power, we could then change the arrangement such that one or a pair of producers gain more from the additional profits than the other players. However, no change in the percentage share of the additional profits could induce one of the pairs of producers to agree without the other pair being worse off. Consequently, in this real-world case, the competitive outcome will always be marginally preferable for at least one of the pairs.

The preference for horizontal collusion depends primarily on the level of asymmetry in potential demand (catchment areas) and the producers' cost functions. Ryanair is an ultra-low-cost carrier whose cost per passenger is 60% that of Easyjet, and the population of the Stansted monopoly catchment area is around double that of Luton. Therefore, the $(S, R)$ pair has less interest in horizontal collusion and prefers the non-collaborative option. We note that the case where both airport–airline pairs vertically collude without airport-airport horizontal collusion (scenario (iv)) also offers the preferable outcome for passengers. This equilibrium serves the highest demand with the lowest airfares, thus increasing consumer surplus. Specifically, the airport cross-subsidizes the two sides of the market by lowering airline charges in exchange for increased passenger numbers, which, in turn, increases commercial revenues and overall airport profits. This would suggest that the UK Civil Aviation Authority was correct to force ownership separation of the airports and remove price regulation at Luton and Stansted.

To confirm our results, we compare the outcome of the demand and the decision variables to the case study, as depicted in Table 9. The catchment area game equilibrium, scenario (iv), is the most similar to the real-world outcome, in terms of passenger volumes and airfares. The model results over-estimate the airport charges, probably due to excess airport market power over the airlines in the catchment area game which assumes a single airline operating per airport. A further confirmation of the likelihood of vertical collusion may be drawn from the fact that the posted charge per passenger at Stansted is £14 to 15 for a Boeing 737–800[6] whereas the financial report suggests a payment of £6 per passenger (Table 9).

## 4. Conclusions and future directions

The methodology presented in this research includes new algorithms to assess ICT data in order to create demand matrices based on geographical regions. Subsequently, we perform regressions over the ICT data in order to explain the drivers of demand, shedding light on passenger behavior. Finally, we formalize a catchment area game that analyzes potential collusion between multiple service providers, which helps to explain producer behavior.

One of the overarching aims of this research is to assess the potential insights that could be drawn from analyzing information and communication technology (ICT) data from mobile phone applications. The passenger demand, which is geographically based, is created from as series of algorithms that characterize mobility patterns at different spatial scales. The innovative methods extract from mobile phone records the aggregated trajectories of users traveling to the airports from home or work and determine whether they are local residents, tourists or airport workers. Based on this information, catchment area maps are visualized and market shares are estimated relatively accurately. However, conflicting signs regarding competition remain unanswered. Whilst London Heathrow

[6] Stansted's conditions of use, including airport charges, 1 April 2017 to 31 March 2018. https://live-webadmin-media.s3.amazonaws.com/media/3277/stal-conditions-of-use-2017-18.pdf.

draws passengers from across the UK, acts as a hub for British Airways and serves inter-continental demand, most of the demand for the remaining airports draws from local catchment areas. Yet, the same final destinations are offered from multiple airports, which could be deemed a sign of a competitive market. Consequently, we develop an additional model for analyzing competition based on the insights collected.

We identify the importance of access and egress times to the airports together with additional socio-economic demand drivers. The regression results suggest that airport choice is partly determined by the road network and the forms of public transportation available. The most important of these is the existence of railway, metro and bus services, which influence the size of the demand from each area.

We also assess competition with respect to pricing strategies between two London airports with clearly identified overlapping catchment areas. The catchment area game analyzes both horizontal and vertical collusion between and across airports and airlines using both analytics and numerics. Whilst it seems that airports and airlines are likely to vertically collude, horizontal collusion between airports is much harder to determine. The symmetric, theoretical equilibrium solution outcome suggests that airports will horizontally collude. However, the real-world solution outcome between two airports of different sizes does not indicate an incentive for collusion. Our formulation, in conjunction with access to high-resolution consumer behavior data, should enable regulators to analyze catchment areas as a means to gain insight as to competition between producers at multiple levels, e.g., the service level or the pricing level. It would also help to ascertain the likelihood of collusion and the need for corrective regulatory actions. Most importantly, our formulation shows that the theoretical analysis may not always accurately predict the likely outcome. The data-driven analysis acts as a decision support tool, solving the conflict between the visualization of the ICT data and the likelihood of competition or collusion between producers.

Future research could apply this type of methodology to other markets, such as supermarkets, which connect wholesalers with consumers, or shopping malls, which connect retailers with consumers. The problem of information asymmetry between private companies and government regulators is an on-going issue and the use of anonymized big datasets may help the regulators to create a fair playing ground and encourage competitive equilibrium outcomes. The same methodology could also be directly applied, *mutatis mutandis*, to any similar markets in which passengers travel to service-provider facilities. It could also be extended to online markets such as Amazon, Alibaba or eBay, where goods are sent directly to the consumers. In such cases, the catchment areas are dictated by the cost of the goods and the shipping times, which depend on complex logistics involving a network of warehouses. In general, the catchment area game could be embedded not in physical space, but rather in consumer decision space such that accessibility and costs are balanced. This would extend the application to virtual markets such as social media connecting advertisers and users (Facebook/Instagram and Google/YouTube), or digital distribution services that provide downloadable content.

An interesting future direction would be to test whether the ICT data accurately estimates not only trip times but also the ground access mode choice. To analyze the data for this purpose requires the identification of a set of 'gateways', namely locations visited along the trip that enable predictions as to the mode chosen. For short trips, the data would need to be collected consistently, with small inter-event times, such that the trip through a gateway is identified reasonably accurately. Finally, we note that big datasets complement, rather than substitute, consumer surveys. Due to the requirements to anonymize the data, we could not directly connect passengers with socio-economic information, such as age, gender and income. This impeded our ability to directly assess specific elasticities. As stated in Harvey (1987), who investigated the multi-airport San Francisco Bay area using a passenger survey, it is important to pay attention to ground access in planning for multi-airport systems, but it is difficult to predict airport utilization without information about market-specific airline schedules. Furthermore, Lieshout (2012) found that Schiphol airport's catchment area changes according to the route served. In conclusion, combining ICT data with consumer surveys in order to approximately identify the final destination of the traveler would further strengthen the results of the catchment area game.

## CRediT authorship contribution statement

**Nicole Adler:** Developed the economic models and undertook the formal economic analysis, Writing – original draft. **Amir Brudner:** Developed the economic models and undertook the formal economic analysis, Writing – original draft. **Riccardo Gallotti:** Developed the algorithms and methodology to create the datasets, Writing – original draft. **Filippo Privitera:** Provided the raw GIS data. **José J. Ramasco:** Developed the algorithms and methodology to create the datasets, Writing – original draft.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The anonymized trajectory data is proprietary and access must be requested from Cuebiq. The rest of the aggregated data and software for analyses are available in the main text, supplementary materials and links provided.

## Acknowledgments

We would like to thank two anonymous reviewers and the Associate Editor at Transportation Research part B: Methodological for very useful comments and suggestions on earlier versions of this paper. Any remaining errors are our own.

We would like to thank participants in SIDs 2018, ATARD European COST project, ORSIS 2018, Euro 2019, ITEA 2019 and GARS 2020 for helpful discussions.

## Funding

This project was supported by the European Union Horizon 2020 BigData4ATM grant (699260) funded by SESAR. José J. Ramasco also acknowledges funding from MCIN/AEI/10.13039/501100011033/FEDER/EU under project APASOS (PID2021-122256NB-C22) and from MCIN/AEI/10.13039/501100011033 under the María de Maeztu Program for units of Excellence in R&D grant MDM-2017-0711. Nicole Adler thanks the Asper Center for Entrepreneurship at the Hebrew University Business School for partial funding of this research.

## Appendix A. Analytical solutions of the catchment area game cases

We assume throughout this paper that the passenger's willingness to pay to travel is higher than the cost of providing the service.

**Assumption 1.** $V > (\alpha + c)$.

### A.1. Competitive scenario (i)

Assuming no collusion, the airport and airline profits are given by Eqs. (9) and (10), respectively. We solve the maximization problem as a two-stage game where, in the first stage, airports set their charges ($p_L$ and $p_S$), and in the second stage, airlines set airfares ($p_E$ and $p_R$) and airports set the prices of amenities ($k_L$ and $k_S$). The solution procedure starts with the second stage and subsequently estimates the results of the first stage. In the symmetric case, we denote $V_E = V_R = V$, $c_E = c_R = c$, $\beta_L = \beta_S = \beta_j$ and $\alpha$ and $\delta$ replace $\alpha_L$, $\alpha_S$ and $\delta_L$, $\delta_S$ respectively. Therefore, the airfares are given by:

$$p_E^{(i)} = p_R^{(i)} = p_i^{(i)} = \frac{(dt + 2V)(51r - 208\beta_j t) - 204\beta_j t(c + \alpha)}{102r - 620\beta_j t}. \tag{A.1}$$

The airport amenities prices are:

$$k_L^{(i)} = k_S^{(i)} = k_j^{(i)} = \frac{51r(-2\alpha - 2c + dt + 2V)}{620\beta_j t - 102r}, \tag{A.2}$$

and the demand in the monopolistic area is given by:

$$y_L^{m,(i)} = y_S^{m,(i)} = y_j^{m,(i)} = \frac{r\left(204\beta_j(V - \alpha - c) + 51dr - 208\beta_j dt\right)}{620\beta_j t - 102r}. \tag{A.3}$$

The airport charge is:

$$p_L^{(i)} = p_S^{(i)} = p_j^{(i)} = \frac{c(280\beta_j t - 102r) + (51r - 140\beta_j t)(dt + 2V) - 340\alpha\beta_j t}{102r - 620\beta_j t}, \tag{A.4}$$

The airlines and airports profits are given by:

$$\pi_E^{(i)} = \pi_R^{(i)} = \pi_i^{(i)} = \frac{3468\beta_j^2 rt(2V - 2\alpha - 2c + dt)^2}{(102r - 620\beta_j t)^2}, \tag{A.5}$$

$$\pi_L^{(i)} = \pi_S^{(i)} = \pi_j^{(i)} = \frac{51\beta_j r(280\beta_j t - 51r)(-2\alpha - 2c + dt + 2V)^2}{2(102r - 620\beta_j t)^2}. \tag{A.6}$$

### A.2. Airport-airline vertical collusion scenario (iv)

Under the assumption that airports collude with their hubbing airline, the profits of the joint airport–airline combination ($\pi_E + \pi_L$ and $\pi_R + \pi_S$) are maximized in a single stage with respect to airfare and the prices of amenities. The airfares are given by:

$$p_E^{(iv)} = p_R^{(iv)} = p_i^{(iv)} = \frac{(dt + 2V)(51r - 70\beta_j t) - 210\beta_j t(c + \alpha)}{102r - 350\beta_j t}. \tag{A.7}$$

The excess profit is then equally distributed between the airport and the airline according to the Nash bargaining solution (Nash, 1950). The airport amenities prices are:

$$k_L^{(iv)} = k_S^{(iv)} = k_j^{(iv)} = \frac{51r(-2\alpha - 2c + dt + 2V)}{350\beta_j t - 102r}, \tag{A.8}$$

and the demand in the monopolistic area is given by:

$$y_L^{m,(iv)} = y_S^{m,(iv)} = y_j^{m,(iv)} = \frac{r\left(210\beta_j(V - \alpha - c) + 51dr - 70\beta_j dt\right)}{350\beta_j t - 102r}. \tag{A.9}$$

In a vertical case scenario, the charges and profit are set through the Nash bargaining game. Therefore we present only the combined profit:

$$\pi_{E+L}^{(iv)} = \pi_{R+S}^{(iv)} = \pi_{i+j}^{(iv)} = \frac{3\beta_j r(2450\beta_j t - 867r)(2V - 2\alpha - 2c + dt)^2}{2(102r - 350\beta_j t)^2}, \tag{A.10}$$

### A.3. Airport horizontal collusion scenario (v)

In the first stage, airports maximize their joint profits ($\pi_L + \pi_S$) with respect to charges. In the second stage, the airlines separately maximize their profit with respect to the airfare and the airports jointly maximize their profits with respect to the pricing of amenities. The optimal airfares in this scenario are:

$$p_E^{(v)} = p_R^{(v)} = p_i^{(v)} = \frac{(dt + 2V)(51r - 238\beta_j t) - 204\beta_j t(c + \alpha)}{102r - 680\beta_j t}. \tag{A.11}$$

The airport amenities prices are:

$$k_L^{(v)} = k_S^{(v)} = k_j^{(v)} = \frac{51r(2V - 2\alpha - 2c + dt)}{680\beta_j t - 102r}, \tag{A.12}$$

and the demand in the monopolistic area is given by:

$$y_L^{m,(v)} = y_S^{m,(v)} = y_j^{m,(v)} = \frac{r\left(204\beta_j(V - \alpha - c) + 51dr - 238\beta_j dt\right)}{680\beta_j t - 102r}. \tag{A.13}$$

The airport charge under case (v) is:

$$p_L^{(v)} = p_S^{(v)} = p_j^{(v)} = \frac{c(340\beta_j t - 102r) + (51r - 170\beta_j t)(dt + 2V) - 340\alpha\beta_j t}{102r - 680\beta_j t}, \tag{A.14}$$

The airlines and airports profits are given by:

$$\pi_E^{(v)} = \pi_R^{(v)} = \pi_i^{(v)} = \frac{204\beta_j^2 rt(2V - 2\alpha - 2c + dt)^2}{(102r - 680\beta_j t)^2}, \tag{A.15}$$

$$\pi_L^{(v)} = \pi_S^{(v)} = \pi_j^{(v)} = \frac{25.5\beta_j r(2V - 2\alpha - 2c + dt)^2}{680\beta_j t - 102r} \tag{A.16}$$

### A.4. Full collusion scenario (viii)

Under full collusion, the single profit function ($\pi_E + \pi_L + \pi_R + \pi_S$) is maximized with respect to airfares and prices of amenities jointly. The optimal airfares are:

$$p_E^{(viii)} = p_R^{(viii)} = p_i^{(viii)} = \frac{(dt + 2V)(51r - 102\beta_j t) - 204\beta_j t(c + \alpha)}{102r - 408\beta_j t}. \tag{A.17}$$

The airport amenities prices are:

$$k_L^{(viii)} = k_S^{(viii)} = k_j^{(viii)} = \frac{51r(2V - 2\alpha - 2c + dt)}{408\beta_j t - 102r}, \tag{A.18}$$

and the demand in the monopolistic area is given by:

$$y_L^{m,(viii)} = y_S^{m,(viii)} = y_j^{m,(viii)} = \frac{r\left(204\beta_j(V - \alpha - c) + 51dr - 102\beta_j dt\right)}{408\beta_j t - 102r}. \tag{A.19}$$

In a full collusion case scenario which includes vertical collusion, the charges and profit are set through the Nash bargaining game. Therefore we present only the combined profit:

$$\pi_{E+L}^{(viii)} = \pi_{R+S}^{(viii)} = \pi_{i+j}^{(viii)} = \frac{25.5\beta_j r\,(2V - 2\alpha - 2c + dt)^2}{408\beta_j t - 102r}, \tag{A.20}$$

Based on Eqs. (A.1)–(A.19), we conclude with the following corollary, focusing on the access cost, $t$, which is impacted by ground-access policy. Fast and convenient access modes, for example, will reduce the value of travel time and, therefore, the general access cost.

**Corollary 1.** *To ensure non-negative demand for an airport in the monopoly and the competitive areas then* $t > \frac{102r}{350\beta_j}$*. As defined in (A.2), (A.8), (A.12) and (A.18), the denominator of the demand function must be positive. The equation that sets the lowest possible limit on* $t$ *is Eq. (A.8),* $350\beta_j t - 102r \geq 0$*.*

## Appendix B. Ranking the analytical results

### B.1. Analyzing airline and airport profits

**Proposition 1.** *The individual profits of an airline (airport) under horizontal collusion (scenario (v)) are lower (higher) than the profits under full competition (scenario (i)).*

$$\pi_j^{(v)} \geq \pi_j^{(i)} \quad \text{and} \quad \pi_i^{(v)} \leq \pi_i^{(i)}.$$

**Proof of Proposition 1.** Comparing profits in scenario (i) to those in scenario (v):

$$\pi_i^{(i)} - \pi_i^{(v)} = \frac{360\beta_j^3 r t^2 (325\beta_j t - 51r)(-2\alpha - 2c + dt + 2V)^2}{(51r - 310\beta_j t)^2 (3r - 20\beta_j t)^2} \tag{B.1}$$

$$\pi_j^{(v)} - \pi_j^{(i)} = \frac{675\beta_j^3 r t^2 (-2\alpha - 2c + dt + 2V)^2}{(51r - 310\beta_j t)^2 (20\beta_j t - 3r)} \tag{B.2}$$

Eqs. (B.1) and (B.2) are positive under Corollary 1. □

For scenarios that include vertical collusion we propose:

**Proposition 2.** *The aggregate profits under vertical collusion (scenario iv) are lower than those under full collusion (scenario viii).*

$$\pi_{i+j}^{(iv)} \leq \pi_{i+j}^{(viii)}$$

**Proof of Proposition 2.** Comparing aggregate profits in scenario (iv) to those in scenario (viii), we find:

$$\pi_{i+j}^{(viii)} - \pi_{i+j}^{(iv)} = \frac{\beta_j^2 r t (96r - 1225\beta_j t)(-2\alpha - 2c + dt + 2V)^2}{4(51r - 175\beta_j t)^2 (r - 4\beta_j t)} \tag{B.3}$$

which is positive under Corollary 1. □

**Proposition 3.** *If* $t \geq 0.434\frac{r}{\beta_j} > \frac{102r}{350\beta_j}$*, the aggregate profits under vertical collusion (scenario iv) are higher than the aggregate profits under full competition (scenario i).*

$$\pi_{i+j}^{(iv)} \geq \pi_i^{(i)} + \pi_j^{(i)}, \quad \text{if} \quad t \geq 0.434\frac{r}{\beta_j}$$

**Proof of Proposition 3.** Comparing aggregate profits in scenario (iv) to those in scenario (i):

$$\pi_{i+j}^{(iv)} - \left(\pi_i^{(i)} + \pi_j^{(i)}\right) = \frac{3\beta_j^2 r t \left(18865000\beta_j^2 t^2 - 8000625\beta_j r t - 83232r^2\right)(-2\alpha - 2c + dt + 2V)^2}{4(51r - 310\beta_j t)^2 (51r - 175\beta_j t)^2} \tag{B.4}$$

which is positive under the condition that $t \geq 0.434\frac{r}{\beta_j}$ □

### B.2. Ranking fares analytically

To understand the relationship between the airfares across scenarios, we search for a threshold that permits rankings. Assuming demand is strictly non-negative, we propose the following:

**Proposition 4.** *$p_i^{(iv)} \leq p_i^{(viii)} \leq p_i^{(i)} \leq p_i^{(v)}$, i.e., the airfare is lowest under vertical collusion, followed by full collusion, full competition, and horizontal collusion.*

**Proof of Proposition 4.** Comparing scenarios (i) and (iv):

$$p_i^{(i)} - p_i^{(iv)} = \frac{3\beta_j t(4900\beta_j t - 51r)(-2\alpha - 2c + dt + 2V)}{2(51r - 310\beta_j t)(51r - 175\beta_j t)}. \tag{B.5}$$

The conditions required for Eq. (B.5) to be positive are:

1. $V > (\alpha + c)$
2. $t > \frac{51r}{175\beta_j}$ or ($t < \frac{51r}{310\beta_j}$ and $t > \frac{51r}{4900\beta_j}$)

Both conditions are met under Assumption 1 and Corollary 1.

Comparing scenarios (iv) and (v):

$$p_i^{(v)} - p_i^{(iv)} = \frac{3\beta_j t(5950\beta_j t - 51r)(-2\alpha - 2c + dt + 2V)}{2(51r - 340\beta_j t)(51r - 175\beta_j t)}. \tag{B.6}$$

The conditions required for Eq. (B.6) to be positive are:

1. $V > (\alpha + c)$
2. $t > \frac{51r}{175\beta_j}$ or ($t < \frac{51r}{340\beta_j}$ and $t > \frac{51r}{5950\beta_j}$)

Both conditions are met under Assumption 1 and Corollary 1.

Comparing scenarios (iv) and (viii):

$$p_i^{(viii)} - p_i^{(iv)} = \frac{3\beta_j t(3570\beta_j t - 51r)(-2\alpha - 2c + dt + 2V)}{2(51r - 204\beta_j t)(51r - 175\beta_j t)}. \tag{B.7}$$

The conditions required for Eq. (B.7) to be positive are:

1. $V > (\alpha + c)$
2. $t > \frac{51r}{175\beta_j}$ or ($t < \frac{51r}{204\beta_j}$ and $t > \frac{51r}{3570\beta_j}$)

Both conditions are met under Assumption 1 and Corollary 1.

Comparing scenarios (i) and (viii):

$$p_i^{(i)} - p_i^{(viii)} = \frac{5406\beta_j^2 t^2 (-2\alpha - 2c + dt + 2V)}{2(51r - 204\beta_j t)(51r - 310\beta_j t)}. \tag{B.8}$$

The conditions for Eq. (B.8) to be positive are:

1. $V > (\alpha + c)$
2. $t > \frac{51r}{204\beta_j}$ or $t < \frac{51r}{310\beta_j}$

Both conditions are met under Assumption 1 and Corollary 1. □

### *B.3. Ranking the prices of airport amenities analytically*

The ranking across scenarios of the price of amenities at the airports is in reverse order to that of the airfares:

**Proposition 5.** $k_j^{(v)} \leq k_j^{(i)} \leq k_j^{(viii)} \leq k_j^{(iv)}$

**Proof of Proposition 5.** The numerators of Eqs. (A.2), (A.8), (A.12) and (A.18) are identical and, given Assumption 1, are also positive. Following Corollary 1, the denominators are also positive and the ranking is immediate. □

## Appendix C. Regression results

The two regression results depicted visually in Fig. 5 are presented numerically in Table C.10.

**Table C.10**
Airport demand regressions.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | MSOA to airport demand (0,1) *logistic* | | ln(MSOA to airport demand) *OLS* | |
| | (1) | (2) | (3) | (4) |
| Constant | −31.833*** (0.890) | −31.133*** (0.899) | −164.185*** (3.693) | −155.785*** (3.707) |
| Population size | 0.134*** (0.007) | 0.133*** (0.007) | 0.606*** (0.031) | 0.644*** (0.032) |
| Income | 0.201*** (0.009) | 0.195*** (0.009) | 1.281*** (0.034) | 1.147*** (0.035) |
| Air traffic movements | 0.117*** (0.002) | 0.117*** (0.002) | 0.474*** (0.009) | 0.478*** (0.009) |
| Overlapping area | −0.481*** (0.047) | −0.486*** (0.048) | −0.246 (0.170) | −0.407** (0.169) |
| Closest alternative (h) | 0.055*** (0.004) | 0.057*** (0.004) | 0.164*** (0.012) | 0.180*** (0.012) |
| Access time (h) | −0.378*** (0.006) | −0.378*** (0.006) | −1.076*** (0.013) | −1.068*** (0.013) |
| Car ownership | −0.081*** (0.004) | −0.070*** (0.006) | −0.859*** (0.016) | −0.692*** (0.022) |
| Number of bus stops | | −0.009*** (0.003) | | −0.088*** (0.010) |
| Has coach station | | 0.333*** (0.038) | | 0.759*** (0.165) |
| Has rail station | | 0.127*** (0.031) | | 1.011*** (0.129) |
| Has metro station | | 0.172*** (0.065) | | 2.089*** (0.221) |
| Observations | 43,197 | 43,167 | 15,981 | 15,956 |
| Pseudo $R^2$ | 0.398 | 0.400 | | |
| $R^2$ | | | 0.542 | 0.550 |
| Adjusted $R^2$ | | | 0.542 | 0.549 |

*Note:* *p < 0.1; **p < 0.05; ***p < 0.01.

## References


Adler, N., Liebert, V., 2014. Joint impact of competition, ownership form and economic regulation on airport performance and pricing. Transp. Res. A 64, 92–109.

Andreyeva, T., Long, M.W., Brownell, K.D., 2010. The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. Am J Public Health 100, 216–222.

Balcan, D., Colizza, V., Gonçalves, H., Ramasco, J.J., Vespignani, A., 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc. Natl. Acad. Sci. USA 106, 21484–21489.

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C.R., Lenormand, M., Louail, T., Menezes, R., Ramasco, J.J., Simini, F., Tomasini, M., 2018. Human mobility: Models and applications. Phys. Rep. 734, 1–74.

Barbot, C., 2009. Airport and airlines competition: Incentives for vertical collusion. Transp. Res. B 43, 952–965.

Bassolas, A., Lenormand, M., Tugores, A., Gonçalves, J.J., 2016. Touristic site attractiveness seen through twitter. EPJ Data Sci. 5, 12.

Bottasso, A., Conti, M., 2012. The cost structure of the uk airport industry. J. Transp. Econ. Policy (JTEP) 46, 313–332.

CAA, 2016. Caa passenger survey report 2016. URL: https://www.caa.co.uk/uploadedFiles/CAA/Content/Standard_Content/Data_and_analysis/Datasets/Passenger_survey/CAA%20Passenger%20Survey%20Report%202016.pdf.

CAA, 2017. Annual airport data 2017. URL: https://www.caa.co.uk/data-and-analysis/uk-aviation-market/airports/uk-airport-data/uk-airport-data-2017/annual-2017/.

Del Chiappa, G., Martin, J.C., Roman, C., 2016. Service quality of airports' food and beverage retailers, a fuzzy approach. J. Air Transp. Manage. 53, 105–113.

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F.R., Gaughan, A.E., Blondel, V.D., Tatem, A.J., 2014. Dynamic population mapping using mobile phone data. Proc. Natl. Acad. Sci. USA 111, 15888–15893.

Fletcher, D., MacKenzie, D., Villouta, E., 2005. Modelling skewed data with many zeros: a simple approach combining ordinary and logistic regression. Environ. Ecol. Stat. 12, 45–54.

Florez, M.A., Jiang, S., Li, R., Mojica, C.H., Rios, R.A., Gonzáles, M., 2018. Measuring the impact of economic well being in commuting networks-a case study of Bogota, Colombia. In: Transportation Research Board 96th Annual Meeting.
Fröhlich, K., Niemeier, H.M., 2011. The importance of spatial economics for assessing airport competition. J. Air Transp. Manage. 17, 44–48.
Fuellhart, K., 2007. Airport catchment and leakage in a multi-airport region: The case of harrisburg international. J. Transp. Geogr. 15, 231–244.
Gallotti, R., Barthelemy, M., 2014. Anatomy and efficiency of urban multimodal mobility. Sci. Rep. 4.
Gallotti, R., Barthelemy, M., 2015. The multilayer temporal network of public transport in great britain. Sci. Data 2.
Gallotti, R., Bazzani, A., Rambaldi, S., 2012. Towards a statistical physics of human mobility. Internat. J. Modern Phys. C 23, 1250061.
Gallotti, R., Bazzani, A., Rambaldi, S., Barthelemy, M., 2016. A stochastic model of randomly accelerated walkers for human mobility. Nature Commun. 7, 12600.
Gaudet, G., Long, N.Van., 1996. Vertical integration, foreclosure, and profits in the presence of double marginalization. J. Econ. Manage. Strategy 5, 409–432.
Gillen, D., 2011. The evolution of airport ownership and governance. J. Air Transp. Manage. 17, 3–13.
Harvey, G., 1987. Airport choice in a multiple airport region. Transp. Res. A 21, 439–449.
Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., Ratti, C., 2014. Geo-located twitter as proxy for global mobility patterns. Cartogr. Geogr. Inf. Sci. 41, 260–271.
Hosni, H., Vulpiani, A., 2017. Forecasting in light of big data. Philos. Technol. 1–13.
Iqbal, M.S., Choudhury, C.F., Wang, P., González, M.C., 2014. Development of origin–destination matrices using mobile phone call data. Transp. Res. C 40, 63–74.
Khan, N., Yaqoob, I., Hashem, I.A.T., Inayat, Z., Ali, M., Kamaleldin, W., Alam, M., Shiraz, M., Gani, A., 2014. Big data: survey, technologies, opportunities, and challenges. Sci. World J. 2014.
Lenormand, M., Gonçalves, A., Ramasco, J.J., 2015a. Human diffusion and city influence. J. R. Soc. Interface 12, 20150473.
Lenormand, M., Louail, T., Cantú-Ros, M., Herranz, R., Arias, J.M., Barthelemy, M., San Miguel, M., Ramasco, J.J., 2015b. Influence of sociodemographics on human mobility. Sci. Rep. 5.
Lenormand, M., Picornell, M., Cantú-Ros, A., Louail, T., Herranz, R., Barthelemy, M., Frias-Martinez, E., Ramasco, J.J., 2014. Cross-checking different sources of mobility information. PLoS One 9, e105184.
Lieshout, R., 2012. Measuring the size of an airport's catchment area. J. Transp. Geogr. 25, 27–34.
Lieshout, R., Malighetti, P., Redondi, R., Burghouwt, G., 2016. The competitive landscape of air transport in Europe. J. Transp. Geogr. 50, 68–82.
Lotero, L., Cardillo, A., Hurtado, R., Gómez-Gardeñes, J., 2016a. Several multiplexes in the same city: the role of socioeconomic differences in urban mobility. In: Interconnected Networks. Springer, pp. 149–164.
Lotero, L., Hurtado, R.G., Floría, J., 2016b. Rich do not rise early: spatio-temporal patterns in the mobility networks of different socio-economic classes. R. Soc. Open Sci. 3, 150654.
Merkert, R., Beck, M., 2017. Value of travel time savings and willingness to pay for regional aviation. Transp. Res. A 96, 29–42.
Nash, Jr., J.F., 1950. The bargaining problem. Econometrica 15, 5–162.
Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., Mascolo, C., 2012. A tale of many cities: universal patterns in human urban mobility. PLoS One 7, e37027.
Pels, E., Nijkamp, P., Rietveld, P., 2001. Airport and airline choice in a multiple airport region: an empirical analysis for the san Francisco bay area. Reg. Stud. 35, 1–9.
Pels, E., Nijkamp, P., Rietveld, P., 2003. Access to and competition between airports: a case study for the san Francisco bay area. Transp. Res. A 37, 71–83.
Rochet, J.C., Tirole, J., 2004. Two-Sided Markets: An Overview. Institut d'Economie Industrielle Working Paper.
Simini, F., González, A., Barabási, A.L., 2012. A universal model for mobility and migration patterns. Nature 484, 5.
Starkie, D., 2002. Airport regulation and competition. J. Air Transp. Manage. 8, 63–72.
Sun, X., Wandelt, S., Zhang, A., 2021. Comparative accessibility of chinese airports and high-speed railway stations: A high-resolution, yet scalable framework based on open data. J. Air Transp. Manage. 92, 102014.
Teraji, Y., Morimoto, Y., 2014. Price competition of airports and its effect on the airline network. Econ. Transp. 3, 45–57.
Thelle, M.H., Pedersen, T.T., Harhoff, F., 2012. Airport competition in europe. Copenhagen Econ..
Vespignani, A., 2012. Modelling dynamical processes in complex socio-technical systems. Nat. Phys. 8, 32.
Yuan, J., Zheng, Y., Xie, X., 2012. Discovering regions of different functions in a city using human mobility and pois. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 186–194.