

Key Event Analysis of Customer Behaviour Using Clickstream Data in Airline Market

Jihao Lin¹, Christopher Holland¹, Nikolaos Argyris¹ and Laura Hornbake²

¹Loughborough University, UK

²Swiss International Air Lines, Switzerland

j.lin@lboro.ac.uk

c.p.holland@lboro.ac.uk

n.argyris@lboro.ac.uk

laura.hornbake@swiss.com

Abstract: COVID-19 travel restrictions have significantly affected airline markets. A behavioural analysis is presented of 1 million customers of a major airline company over 3 years, using clickstream data, combined with COVID-19 Government Response Tracker (OxCGRT) data. The main results are that pre-pandemic, the variations in traffic volumes were seasonal and there was a consistent ratio of search (76%), buy (5%) and e-service (19%) activities across the year. The start of the pandemic caused a sharp drop in website traffic, which minimised at 1/3 of normal levels and then gradually increased so that post-covid volumes are now 20% higher, which reflects pent-up demand and consumer confidence. At the start of the pandemic, e-service increased sharply and then dropped to below pre-pandemic levels. The proportion of buyers decreased significantly at the start of the pandemic and then returned to just below pre-pandemic levels. The proportion of searchers remained high throughout the pandemic. Comparing post-pandemic with pre-pandemic, the proportion of searchers remains higher, buyers has dropped, and e-service remains the same. These results show the importance of distinguishing between different types of online activity, defined by three broad phases in the customer lifecycle: search, buy and e-service, which can only be achieved by using raw, clickstream data. Typically, web server analytics software employ aggregate data and it is much more difficult to uncover nuanced changes that are apparent at the granular level of individual clickstream data. The contribution of the research is to demonstrate a novel analytical methodology that exploits detailed clickstream data to characterize the dynamic changes in customer behaviour, which gives important insights into how online markets change in response to exogenous shocks by enabling key event analysis. Using the approaches outlined in this paper, airlines and e-commerce companies can anticipate, measure behavioural changes and be pro-active in their strategic responses to significant external events in the competitive environment that are likely to affect customer behaviour.

Keywords: Airline market, Clickstream data, Web server, Customer lifecycle, Exogenous shocks

1. Introduction

The travel and tourism industries are faced with difficult decisions in times of crisis such as economic downturns and pandemics like COVID-19. Many airlines have suffered unprecedented losses and are under significant financial pressure due to the outbreak of the virus. Throughout the pandemic, decision-makings of airlines are focused on the spread of the virus and the corresponding travel restrictions imposed by the governments (Suk and Kim, 2021).

Unlike the previous crisis times such as September 11 terrorist attacks and the SARS outbreak, the high spread speed and continuous evolving of COVID-19 has impacted the world much longer and make it difficult to assess the risks. The United Nations World Tourism Organisations initially estimated the 1-3% decline in international travel in early March, and quickly updated this figure to 20-30% towards the end of March (Gössling et al, 2020). This shows the difficulty of calculating the potential consequences at the time. The governmental responses to tackle the virus include school and workplace closure, travel restrictions and bans on public gatherings. However, these policies vary from country to country, and affect the travel industry differently.

The popular topics of the studies on the pandemic impact on the airline industry include the customer demand, economic consequences and airport safety measures, however, limited studies have been conducted to investigate the nature and scale of the online customer behaviour shift. We use the clickstream data spanning across three years, from July 2019 to July 2022, to examine the key events of COVID19 such as the initial outbreak, global and local lockdowns and the relaxations of travel restrictions and infer the change of online customer search and buying behaviours. This allows us to have a closer look of how customers behave in different time periods. Previous studies on this topic largely use survey data, or aggregated data (Bulchand-Gidumal & Melián-González, 2021; MacSween and Canziani, 2021). The survey data cannot reflect the rapid change of the situation as we have seen that the government policies to tackle the pandemic could change overnight. The aggregated data does not provide us with the detailed information such as what individual customers browsed and searched, hence, the insight provided are more general. Clickstream data are records

of online sequential events of a website with associated time stamps. It provides fine-grained information about how customers navigate through websites, including where they land on a website, the web pages viewed, time spent on each page etc. Firstly, we examine the website traffic volume against the strictness of the governmental policies and found that there is a negative relationship between the two. We then further segment the customers into searcher, buyer and e-service user, based on their purpose of visit. A longitudinal plot of the market composition across three years is given, the change of the composition is analysed through the lens of key pandemic events.

The paper is structured as follows: a literature review that includes the studies of COVID-19 impact on the airline industry and clickstream data research is given. We then explain our data set, how to process them, and methodology in details. In section 4, we provide our key results and findings, followed by the discussions and conclusions.

2. Literature Review

We conducted the literature review in two aspects, the studies of examining the impact of COVID-19 on the airline industry and the research on using clickstream data to uncover online consumer behaviour.

2.1 Covid-19 and Airline Industry

The sudden demand and supply shocks caused by the pandemic are in almost in every sector, which is also reflected by the shift in the consumer purchase behaviour. The tourism sector is the one that has been affected heavily since many governments impose the restrictions on movements to slow down the spread of the disease.

The drastic decrease in air traffic resulted in 43 airlines ceased or suspended operations in 2020 (CNBC, 2020). The massive decline in demand, coupled with the expected long term recovery period, imposes many challenges for the sector. These difficulties include increased need in faster and more flexible decision-making and forecasting demand during and after COVID-19 (Garrow et al, 2021).

Various case studies have been conducted to assess the pandemic impact on the airline companies and how they responded under this turbulent time. Pereira and Mello (2021) analysed the operational response to the disease outbreak for the main Brazilian airlines and concluded that the company with a better mix of aircraft models is more efficient to minimise the revenue loss. Kim and Sohn (2022) showed that South Korean airlines mitigated the crisis through mergers and acquisitions, concentration on cargo and diversification of business models etc. The study by Ng et al (2022) focuses on the effect of the pandemic on the capacity, frequency, yield of route and competition of Japanese airlines. In the Chinese airline market, the air carriers that are not well funded and whose networks are mainly concentrated on international markets, premium traffic and leisure travel have been impacted heavily by COVID-19 (Warnock-Smith et al, 2021). In Europe, the most common responses by airlines to overcome these challenges were partially suspending operations, temporary groundings of aircrafts, reducing workforce and changes to network and capacity (Budd et al, 2020).

The pandemic also has affected people's travel decision-making. In general, the perceived risk played an important role in travel avoidance and cancellation in the previous infectious disease outbreaks (Cahyanto et al, 2016; Leggat et al, 2010; Rittichainuwat and Chakraborty, 2009), this is further confirmed by Neuburger et al (2021) surveying the travellers in some European regions identifying three groups of tourists based on their perception of COVID-19. It has been studied that the vaccine passport influences tourists' travel intention (Shah et al, 2022), Shin et al. (2021) concluded that the vaccine passports reduced the health risk perception of air travellers, hence, affecting their flight booking intentions. In addition, travellers are more reluctant to book and pay in advance for their services and prefer to buy their tickets near the departure date in the future, especially for airline ticket (Bulchand-Gidumal and Melián-González, 2021). Customers rely on online information search to make decisions, especially in travel and tourism industry, the internet has become the first choice to acquire information and been utilised for marketing tourism services (Sotiriadis, 2017). However, limited research has been conducted so far regarding how consumers search information when planning travels in the era of the pandemic. A study conducted by MacSween and Canziana (2021) examines the information search behaviour in the initial outbreak of COVID-19 and suggests that travel search acted as a source of entertainment during lockdowns. Gallego and Font (2021) used the search and pick data from the metasearch engine Skyscanner to detect the reactivation of tourism markets. We believe this is the first study conducted to look at how online consumer search and buying behaviour change across several key events associated with COVID-19 in the airline industry.

2.2 Clickstream Data Analysis

Clickstream data contains Internet users' footprint, and it plays a fundamental role in studying online consumer search and buying behaviours. Segmenting online customers using clickstream data is of great interest to many researchers. Such study normally involves extracting users' behavioural features from the unstructured clickstream data, e.g., page browsing frequency and variety, search depth, time spent on each page and so on. Moe (2003) uses clickstream data to validate the consumer typology proposed based on in-store navigational patterns. Schellong et al (2016) follows the similar research framework as Moe (2003) and uses off-site clickstream data to segment online customers into buying, browsing, searching and bouncing. Zavali et al. (2021) measures consumers' browsing behaviour using the frequency and intensity of the visit, combined with traffic source and device used, to cluster online shoppers and investigate the revenue generated by each consumer type.

For market competitor analysis, clickstream data is also widely used, as it provides a gateway into examining how consumers choose between different brands. In these type of studies, the user-centric clickstream data, also known as online panel data, is used. This type of data tracks customers activities across different websites, which enables academics and practitioners to measure how a particular website performs compared to their competitors. Wood et al (2005) uses such data to estimate the market share and positioning in the library sector. Holland and Leefmans (2013) measures the online performance of both market leaders and small banks in Mexico by using online panel data. In the airline industry, Holland et al (2020) utilises network analysis to infer the relationships between competitors, and between competitors and search intermediaries.

3. Data and Methodology

This study is based on a collaboration project with a major European airline. Web server logs are used to understand the change of search and buying behaviours across 3 years and the results are analysed in conjunction with the exogenous shocks of the pandemic.

3.1 Data Set

Site-centric clickstream data records detailed interactions between a visitor and a website, ranging from where they come from, i.e. traffic source, the device and browser they use, and the pages visited and the time spent on them etc. In our data, it also contains the customers' search history, such as origin and destination, departure and return date and days to their departure etc. Table 1 shows how our tabular clickstream data are formatted.

Table 1: An Example of Site-Centric Clickstream Data

User ID	Session ID	Date	Page Name	Origin & Destination	Departure	Return	Days to departure
10035	589642	15/08/21	Homepage				
10035	589642	15/08/21	Flight Selection	lhr-fra	23/08/21	25/08/21	8
10036	223688	15/08/21	Item Confirmation	fra-bcn	27/08/21	31/08/21	12
10036	223688	15/08/21	Passenger Details				
10036	223688	15/08/21	Payment Details				
10036	223688	15/08/21	Order Confirmation				

3.2 Clickstream Data Pre-Processing

These clickstream data need to be processed for further analysis. Cooley et al (1999) defines four steps for clickstream data pre-processing, data cleaning, user identification, session identification and path completion. Data cleaning refers to the step of deleting log entries that are irrelevant to our interest, such as graphics and scripts. The modern web analytics software such as Google Analytics deals with this problem automatically and allows users to specify what types of information they want to keep. The websites implement user identification by assigning randomised user ID. The task of session identification is dividing the clickstream data into different chunks of individual sessions. This helps us to know when a user leaves the website and will not be back for a pre-defined threshold, which is normally 30 minutes in practice. The problem of path completion occurs due to

the browser cache function, which are also dealt with by modern web analytics software. After the pre-processing, the data are fed into the web analytics tool to build real-time reports, they can also be fetched to do detailed analysis.

3.3 Imposing Market Segments

We categorise the online customers into three types based on their purpose of visit, which are searchers, buyers, and e-service users, instead of traditionally searching and buying activities. E-service users are often ignored because marketers and practitioners focus on buying outcomes and searching behaviours. However, as our analysis shows, e-service users dominated the traffic during the initial COVID-19 period and continued to play an important part throughout the period of study. Imposing such market segments helps us to better understand the change in dynamics of the customers, which is reflected by the change of composition for these three segments.

3.4 COVID-19 Stringency Index

The government across the world responded different to tackle the spread of the virus by imposing restrictions on travel, closing schools, banning on gatherings. The Oxford COVID-19 Government Response Tracker (OxCGRT) (Hale et al., 2021) measures how tight these restrictions are, by combining indicators that track containment and closure, economic response and health systems. There are four indices have been built, which differentiate each other by using different combinations of indicators. We use the stringency index in this study, which is also widely used in the previous studies in COVID-19 impact on the travel and tourism industry (Chen et al, 2020; Kaczmarek et al, 2021).

4. Results

As stringency index is calculated for each individual country, we separate the customers based on the geographic locations from which they visit the website. Since there are 249 countries/regions in our data set, we focus on analysing the 3 countries that account for the majority of visits, they are Germany, United States and United Kingdom.

4.1 Website Traffic and Stringency Index

The stringency index measures how tight the governments' restrictions are, it is expected that it negatively correlates with the website traffic. We group the dates with the same stringency index and calculate the average website traffic for each country, and plot them in Figure 1, with a, b, c for Germany, United States and United Kingdom, respectively. The vertical axes are masked as they are commercially sensitive. The corresponding linear regression equation and R^2 value is also displayed in the plots.

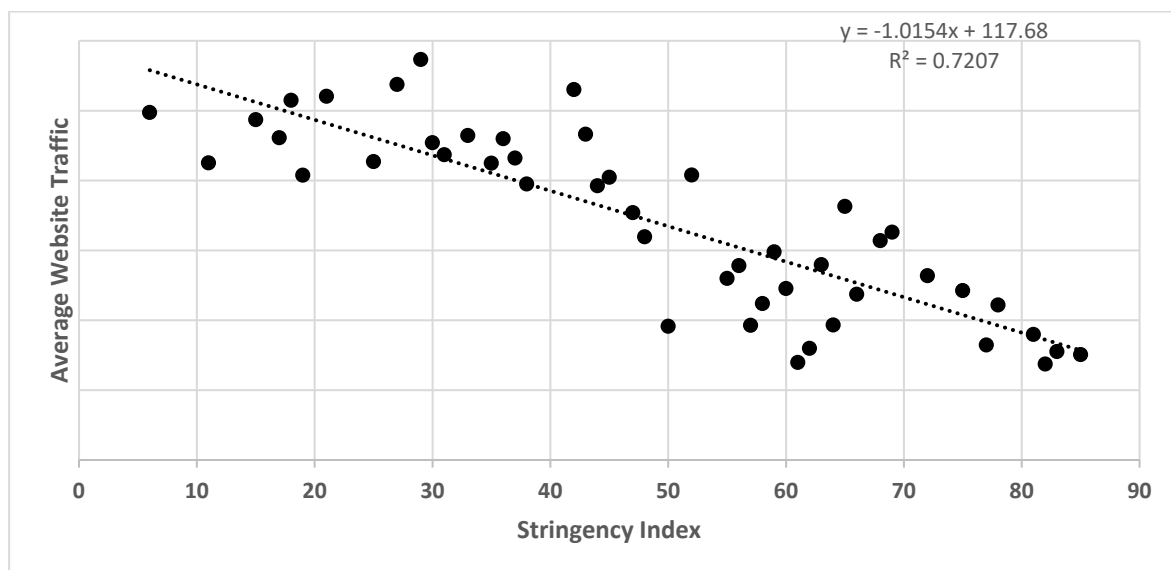


Figure 1a: Average Website Traffic vs Stringency Index – Germany

From Figure 1a, we can clearly see that there is a strong negative relationship between the website traffic and the stringency index for German visitors. The high value of R^2 also confirms the strength of this negative correlation. For visitors from United States, the relationship is not as strong, as we have seen two obvious

outliers, which are the two points having the highest average website traffics and are associated with high stringency index. These two data points are from 15th Feb to 2nd Mar 2022, the period that the restrictions remained high, but the air travel had a strong rebound (IATA, 2022). The UK also displays an inverse trend that as the stringency index increases, the website traffic drops. The three outliers above the regression line also come from Feb to early March 2022.

The study confirmed that the pandemic related restrictions changed how people use different transportation, with airplanes and buses are deemed to be the riskiest transport modes (Barbieri et al, 2021). The existing study in tourism and travel industry also confirm that the pandemic restrictions have a significant negative impact on people's travel (Gupta et al, 2021). Apart from the fact that the restrictive measures discourage the travel intention, tourists also perceive travel risk differently under COVID-19 and reduce their travel plan to destinations (Rahman et al, 2021; Sánchez-Cañizares et al, 2021). This is also reflected in the above that the tighter the restrictions become, the lesser customers are browsing the website. However, the traffic volume does not reveal the customer visiting purpose, in the next subsection, we take a closer look on the pandemic impact on how customers use the website.

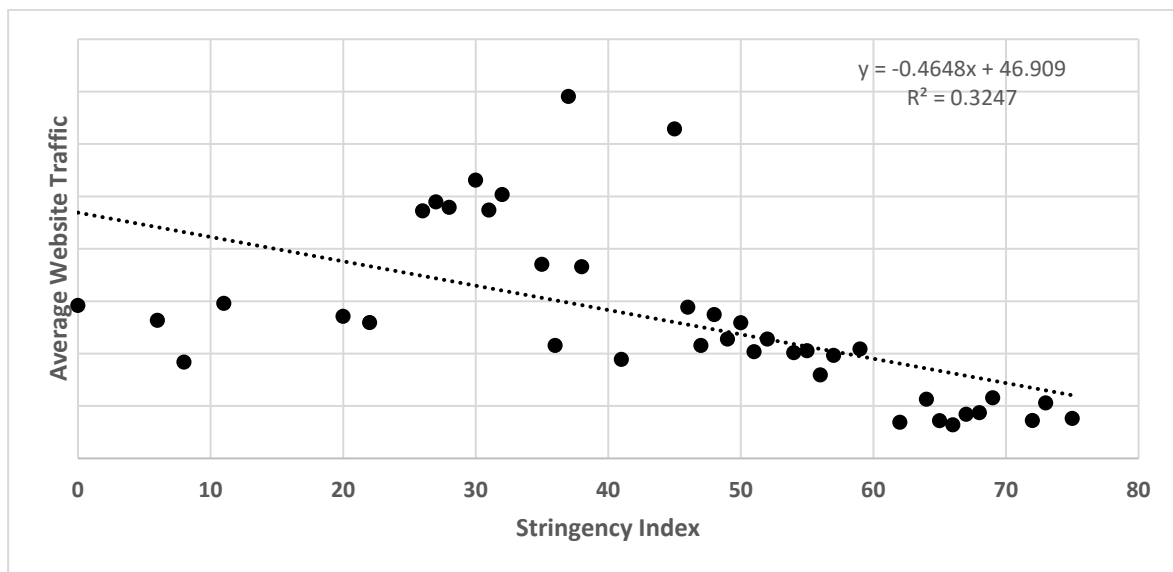


Figure 1b: Average Website Traffic vs Stringency Index – United States

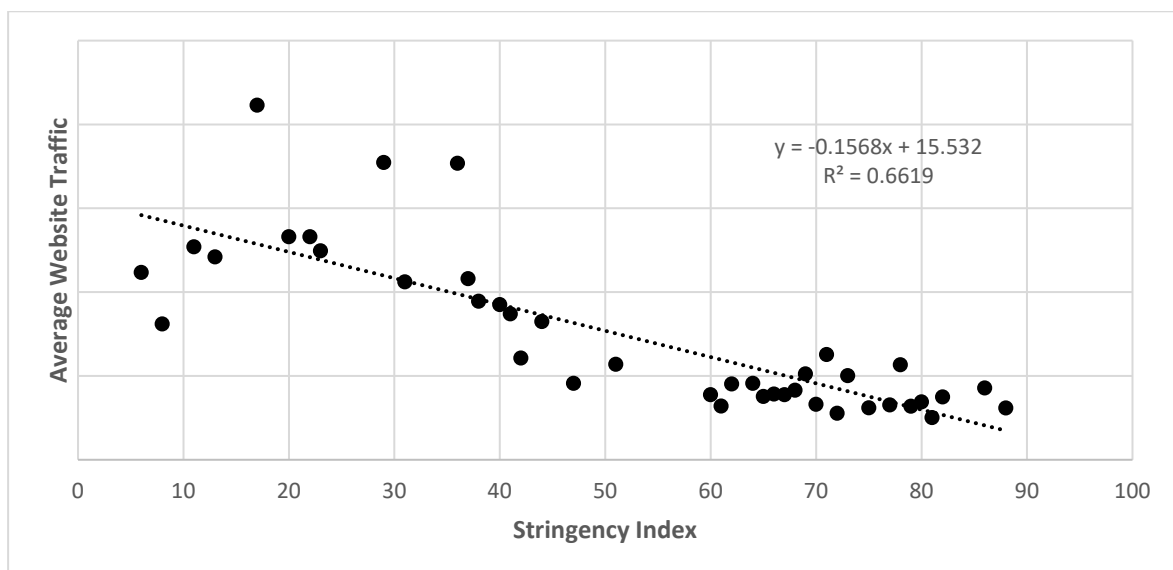


Figure 1c: Average Website Traffic vs Stringency Index – United Kingdom

4.2 Market Composition

The market segmentation helps us to understand the shift of consumer behaviour, which can be observed in the following figures. The Figure 2a-c shows the weekly market composition of searcher, buyer and e-service users across three years for Germany, United States and United Kingdom, alongside the stringency index. From the plots, we can see that whenever there was a significant change in the stringency index, there was drop in both searchers and buyers, and increase in e-service users.

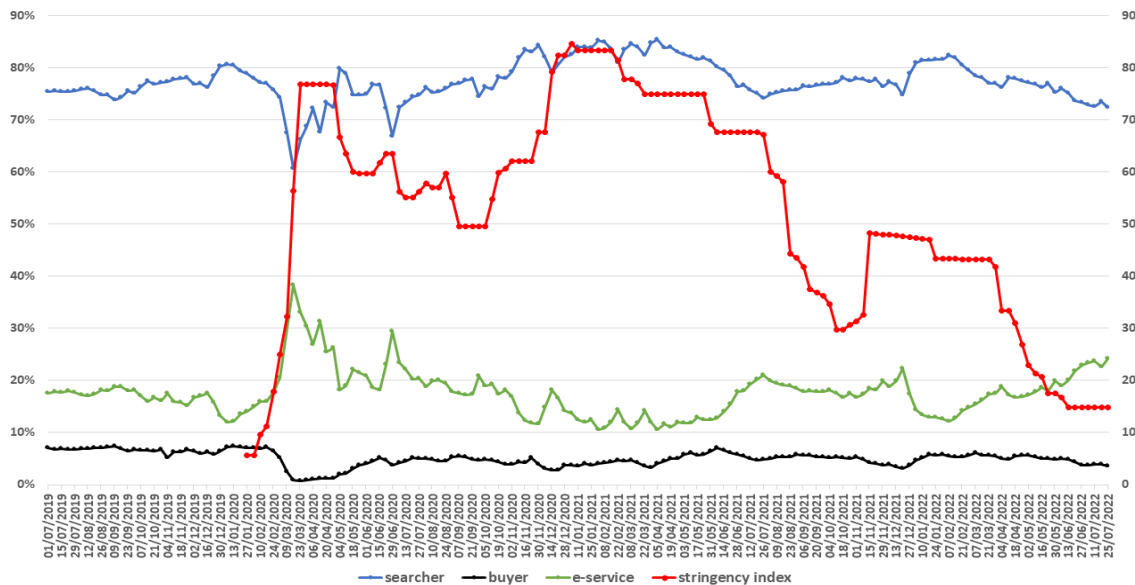


Figure 2a: Market Composition by Week (01/07/2019 - 31/07/2022) – Germany

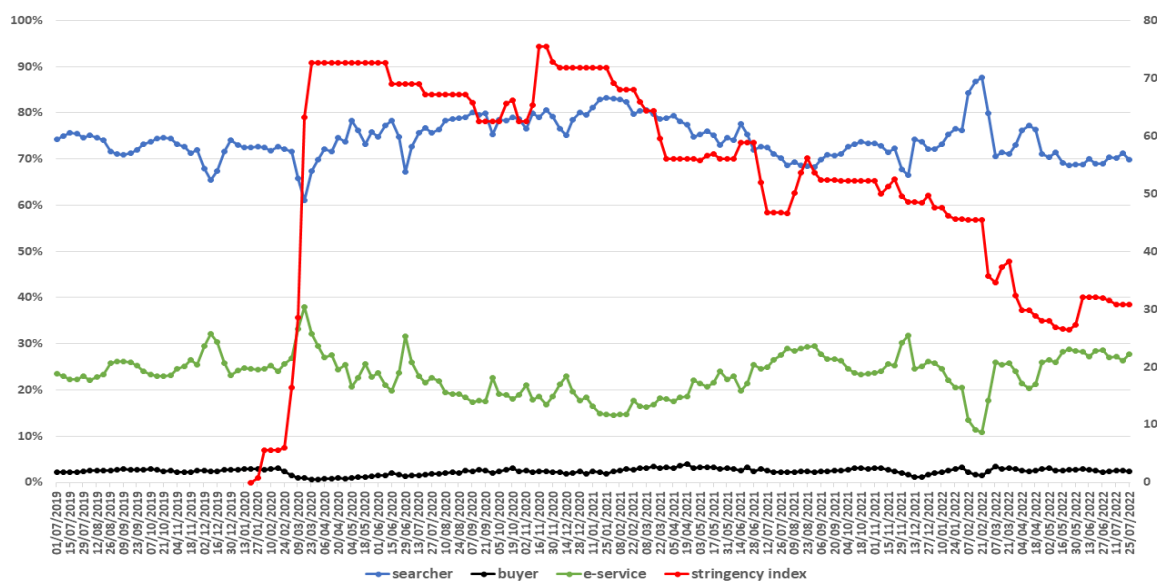


Figure 2b: Market Composition by Week (01/07/2019 - 31/07/2022) – US

The apparent pattern can be observed in March 2020, the early stage of the pandemic, when lockdown and travel restrictions were imposed worldwide. A second abrupt increase of e-service users happened at the end of June 2020, when local lockdowns across Europe took place (Glass, 2020). Starting from 23rd November, which corresponds to the date when Germany announcing to extend and tighten restrictions until at least 20th December, the e-service segment continued to increase and reached the peak in the week of 14th December. There was also an increase in both searcher and buyer composition in January and February 2022, which shows a strong passenger rebound. This is also confirmed by IATA study (IATA, 2022). In summer 2022, travellers faced severe disruption in Europe, with the average delay per flight reaching 5-year high (Eurocontrol, 2022). The waves of cancellations and other service disruptions worldwide due to staff shortage in airports, ground operations and flight crews contributes to the uptick in e-service in this period of time.

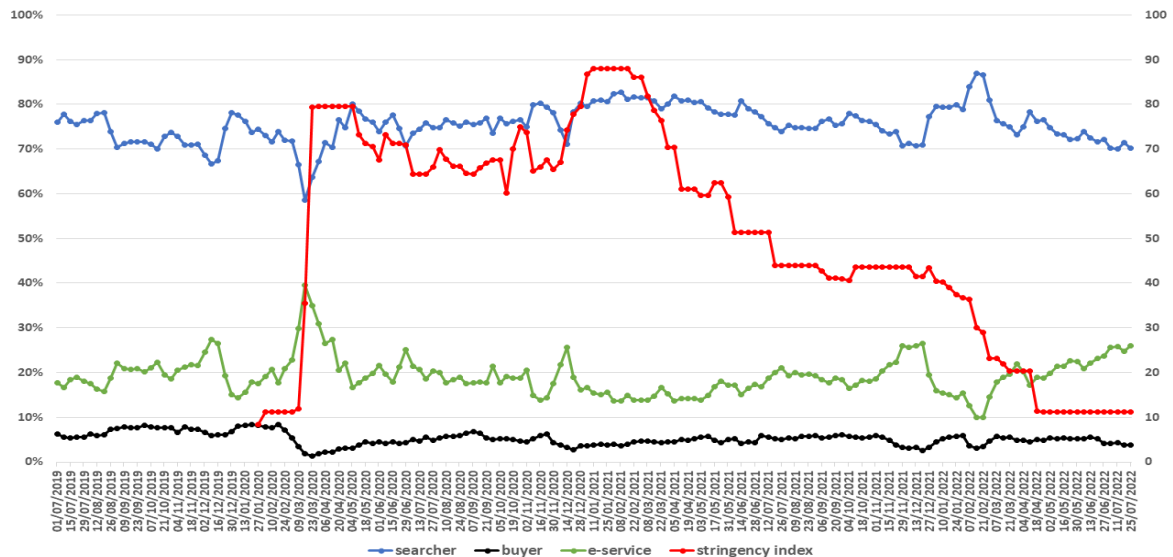


Figure 2c: Market Composition by Week (01/07/2019 - 31/07/2022) – UK

5. Discussion

In this study, the use of Big Data can be understood via the lens of 7Vs, which are volume, velocity, variety, variability, veracity, visualisation and value. The volume aspect is that we have used million sessions across 3 years to enable a detailed longitudinal analysis of the online consumer behaviour. The amount of the clickstream data generated each day in this study shows the high velocity of Big Data. The fact that the huge amount of data can be processed in the same manner easily shows the high velocity. The combination of the clickstream data with COVID-19 stringency index shows the variety of data source we used. Variability, which considers the inconsistency of the data flow, is avoided by using the commercial software such as Google Cloud Platform. The veracity is reflected by using the real data from a major European airline, instead of estimating the result that can cause potential errors. We also provided a good visualisation of our result by plotting the relationship between the stringency index and the website visit, and the market composition longitudinally. Our work shows the Value of using Big Data by showing the market dynamics in times of crisis and providing valuable insights to managers.

Our result shows that the airline website traffic is highly correlated with the stringency index, which measures how tight the governments' policies regarding COVID-19 restrictions. In general, the stricter policy is associated with lower traffic volume. It is without surprise that as more travel restrictions are imposed, people tend not to purchase airline tickets. The deviation between the actual demand and forecast in the flight ticket purchase is huge in the presence of the pandemic, the high correlation between the stringency index and website traffic volume shows the potential of using external data source to help with the demand prediction. The market composition analysis gives a clear picture of how the purpose of customer visit changes through time. We see that the change in the governments' policies leads to more e-service usage, due to people altering their travel plans. This demonstrates the importance of e-service and airline websites need to improve the quality of e-service to better accommodate customers' need.

The previous study conducted by Gallego and Font (2021) utilised the flight search data to show the change of demand in the airline industry, but they failed to use a metric to measure the strictness of governments' policy response to tackle the pandemic and link it with the shift in customer behaviour. Furthermore, rather than just looking at flight search, we provide a clear picture of how the market composition of searcher, buyer and e-service user evolved, which shows the shift in online customer visiting purpose.

6. Conclusions

In this study, we have shown that how we use clickstream data to better understand the changing dynamics of the customer behaviour across time, by linking it with COVID-19. We have done this by mining and analysing the clickstream data of the airline website. It gives managers a powerful tool to closely examine the market composition and support with the decision-making. During the crisis time, it is even more important for organisations to make data-driven decisions, especially in the travel and tourism industry. However, the main

source to develop the tourism statistics remain to be surveys, which is very time and resource consuming and often fails to reflect the rapid change of the situation (Saluveer et al, 2020). We have developed a methodology to inspect the relationship between the pandemic and the search, buying and e-service activities and shown the power of Big Data. The incorporation of the stringency index highlights the importance of using external data sources to uncover customer behaviour. The search queries such as origin-destination pair, days to departure, breadth and depth of the search behaviour can be further explored in future research.

Acknowledgements

The authors would like to thank Laura Hornbake for her help and the collaboration in this project. Further, they want to acknowledge Lufthansa for providing the research data for this paper. Please see <https://www.lufthansa.com/> for further information about Lufthansa. The authors have conducted the analysis and interpretation of the data, and any errors in the paper are the sole responsibility of the authors.

References

- Barbieri, D.M., Lou, B., Passavanti, M., Hui, C., Hoff, I., Lessa, D.A., Sikka, G., Chang, K., Gupta, A., Fang, K. and Banerjee, A. (2021) Impact of COVID-19 pandemic on mobility in ten countries and associated perceived risk for all transport modes. *PloS one*, 16(2), p.e0245886.
- Bawa, A. and Kansal, P. (2008) Cognitive dissonance and the marketing of services: Some Issues. *Journal of Services Research*, 8(2).
- Budd, L., Ison, S. and Adrienne, N. (2020) European airline response to the COVID-19 pandemic—Contraction, consolidation and future considerations for airline business and management. *Research in Transportation Business & Management*, 37, p.100578.
- Bulchand-Gidumal, J. and Melián-González, S. (2021) Post-COVID-19 behavior change in purchase of air tickets. *Annals of Tourism Research*, 87, p.103-129.
- Cahyanto, I., Wiblishauser, M., Pennington-Gray, L. and Schroeder, A. (2016) The dynamics of travel avoidance: The case of Ebola in the US. *Tourism Management Perspectives*, 20, pp.195-203.
- Chen, M.H., Demir, E., García-Gómez, C.D. and Zaremba, A. (2020) The impact of policy responses to COVID-19 on US travel and leisure companies. *Annals of Tourism Research Empirical Insights*, 1(1), p.100003.
- CNBC (2020) Over 40 airlines have failed so far this year — and more are set to come. URL: <https://www.cnbc.com/2020/10/08/over-40-airlines-have-failed-in-2020-so-far-and-more-are-set-to-come.html>
- Cooley, R., Mobasher, B. and Srivastava, J. (1999) Data preparation for mining world wide web browsing patterns. *Knowledge and information systems*, 1, pp.5-32.
- da Silveira Pereira, D. and de Mello, J.C.C.S. (2021) Efficiency evaluation of Brazilian airlines operations considering the Covid-19 outbreak. *Journal of Air Transport Management*, 91, p.101976.
- Eurocontrol (2022) All-Causes Delays to Air Transport in Europe Annual 2022. URL: <https://www.eurocontrol.int/publication/all-causes-delays-air-transport-europe-annual-2022>
- Gallego, I. and Font, X. (2021) Changes in air passenger demand as a result of the COVID-19 crisis: using Big Data to inform tourism policy. *Journal of Sustainable Tourism*, 29(9), pp.1470-1489.
- Garrow, L. and Lurkin, V. (2021) How COVID-19 is impacting and reshaping the airline industry. *Journal of Revenue and Pricing Management*, 20, pp.3-9.
- Glass, D.H. (2020) European and US lockdowns and second waves during the COVID-19 pandemic. *Mathematical biosciences*, 330, p.108472.
- Gössling, S., Scott, D. and Hall, C.M. (2020) Pandemics, tourism and global change: a rapid assessment of COVID-19. *Journal of sustainable tourism*, 29(1), pp.1-20.
- Gupta, V., Cahyanto, I., Sajnani, M. and Shah, C. (2021) Changing dynamics and travel evading: a case of Indian tourists amidst the COVID 19 pandemic. *Journal of Tourism Futures*.
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S. and Tatlow, H. (2021) A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature human behaviour*, 5(4), pp.529-538.
- Holland, C.P. and Leefmans, M.M.G. (2013) Online Consumer Behaviour and Competitor Performance in the Mexican Bank Market. In UKAIS (p. 15).
- Holland, C.P., Thornton, S.C. and Naudé, P. (2020) B2B analytics in the airline market: Harnessing the power of consumer big data. *Industrial Marketing Management*, 86, pp.52-64.
- IATA (2022) Passenger Recovery Accelerates in February. URL: <https://www.iata.org/en/pressroom/2022-releases/2022-04-06-02/> www.iata.org/en/pressroom/2022-releases/2022-04-06-02/
- Kaczmarek, T., Perez, K., Demir, E. and Zaremba, A. (2021) How to survive a pandemic: The corporate resiliency of travel and leisure companies to the COVID-19 outbreak. *Tourism Management*, 84, p.104281.
- Kaffash, S. and Khezrimotlagh, D. (2023) US network and low-cost carriers' performance in response to COVID-19: Strictness of government policies and passengers' panic. *Research in Transportation Business & Management*, 46, p.100835.

- Kim, M. and Sohn, J. (2022) Passenger, airline, and policy responses to the COVID-19 crisis: The case of South Korea. *Journal of air transport management*, 98, p.102144.
- Leggat, P.A., Brown, L.H., Aitken, P. and Speare, R. (2010) Level of concern and precaution taking among Australians regarding travel during pandemic (H1N1) 2009: results from the 2009 Queensland Social Survey. *Journal of travel medicine*, 17(5), pp.291-295.
- MacSween, S. and Canziani, B. (2021) Travel booking intentions and information searching during COVID-19. *International Hospitality Review*, 35(2), pp.208-224.
- Moe, W.W. (2003) Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *Journal of consumer psychology*, 13(1-2), pp.29-39.
- Neuburger, L. and Egger, R. (2021) Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: A case study of the DACH region. *Current issues in tourism*, 24(7), pp.1003-1016.
- Ng, K.T., Fu, X., Hanaoka, S. and Oum, T.H. (2022) Japanese aviation market performance during the COVID-19 pandemic- Analyzing airline yield and competition in the domestic market. *Transport Policy*, 116, pp.237-247.
- Rahman, M.K., Gazi, M.A.I., Bhuiyan, M.A. and Rahaman, M.A. (2021) Effect of Covid-19 pandemic on tourist travel risk and management perceptions. *Plos one*, 16(9), p.e0256486.
- Rittichainuwat, B.N. and Chakraborty, G. (2009) Perceived travel risks regarding terrorism and disease: The case of Thailand. *Tourism management*, 30(3), pp.410-418.
- Saluveer, E., Raun, J., Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., Aasa, A. and Silm, S. (2020) Methodological framework for producing national tourism statistics from mobile positioning data. *Annals of Tourism Research*, 81, p.102895.
- Sánchez-Cañizares, S.M., Cabeza-Ramírez, L.J., Muñoz-Fernández, G. and Fuentes-García, F.J. (2021) Impact of the perceived risk from Covid-19 on intention to travel. *Current Issues in Tourism*, 24(7), pp.970-984.
- Schellong, D., Kemper, J. and Brettel, M. (2016) Clickstream data as a source to uncover consumer shopping types in a large-scale online setting, *European Conference on Information Systems 2016*.
- Shah Alam, S., Masukujjaman, M., Omar, N.A., Mohamed Makhbul, Z.K. and Helmi Ali, M. (2022) Protection Motivation and Travel Intention after the COVID-19 Vaccination: Fear and Risk Perception. *Journal of Quality Assurance in Hospitality & Tourism*, pp.1-27.
- Sotiriadis, M.D. (2017) Sharing tourism experiences in social media: A literature review and a set of suggested business strategies. *International Journal of Contemporary Hospitality Management*.
- Suk, M. and Kim, W. (2021) COVID-19 and the airline industry: crisis management and resilience. *Tourism Review*, 76(4), pp.984-998.
- Warnock-Smith, D., Graham, A., O'Connell, J.F. and Efthymiou, M. (2021) Impact of COVID-19 on air transport passenger markets: Examining evidence from the Chinese market. *Journal of air transport management*, 94, p.102085.
- Zavali, M., Lacka, E. and De Smedt, J. (2021) Shopping hard or hardly shopping: Revealing consumer segments using clickstream data. *IEEE Transactions on Engineering Management*.