# An Analysis of the Impact of Low Cost Airlines on Tourist Stay Duration and Expenditures

Qiu, Wanling and Rudkin, Simon and Sharma, Abhijit

Blue Mountains International Hotel Management School, Torrens University, Suzhou, China, Research Institute of Big Data Analytics, Xi'an Jiaotong-Liverpool University, Suzhou, China, University of Bradford School of Management, Bradford, United Kingdom

14 September 2017

# An Analysis of the Impact of Low Cost Airlines on Tourist Stay Duration and Expenditures

Wanling Qiu*[1], Simon Rudkin†[2], and Abhijit Sharma‡[3]

[1]*Blue Mountains International Hotel Management School, Torrens University, Suzhou, China*
[2]*Research Institute of Big Data Analytics, Xi'an Jiaotong-Liverpool University, Suzhou, China*
[3]*University of Bradford School of Management, Bradford, United Kingdom*

September 14, 2017

**Abstract**

Low cost carriers (budget airlines) have a significant share of the air travel market, but little research has been done to understand the distributional effect of their operation on key tourism indicators such as length of stay and expenditure. Using data on European visitors to the United Kingdom we demonstrate how counterfactual decompositions can inform us of the true impact of mode of travel. Passengers on low cost carriers tend to spend less, particularly at the upper end of the distribution. Budget airline users typically stay longer, though differences in characteristics of observed groups are important to this result. Counterfactual techniques provide additional valuable insights not obtained from conventional econometric models used in the literature. Illustrating an application of the methodology to policy we demonstrate that enabling respondents to extend their stay generates the greatest additional expenditure at the lower end of the distribution. We also show nationality is a significant characteristic, with important impacts across the expenditure distribution.

Keywords: low cost carriers, tourist expenditure, counterfactual decomposition
JEL Classifications: R4, R41

*Email: wanling.qiu@laureate.edu.au
†Email: simonrudkin@outlook.com
‡Email: a.sharma12@bradford.ac.uk

# 1 Introduction

Low cost carriers (LCCs) hold a significant share of the air transportation market. Nevertheless the economic impact of LCCs on travel and tourism is understudied. Previous research on LCCs (e.g. Eugenio-Martin and Inchausti-Sintes (2016) and Ferrer-Rosell and Coenders (2017)) have mainly focused on average effects and little attempt has been made to assess differences in characteristics of people flying with LCCs and traditional carriers (TC) which is a significant omission. This research makes a significant methodological contribution by employing the counterfactual decomposition technique of Chernozhukov et al. (2013) in transportation and tourism for the first time. LCC use is a binary decision making its impact analogous to a treatment, the dustributional study of which is best performed with distributional regressions like Chernozhukov et al. (2013) (Jones et al., 2015). Hence we are able to provide insights additional to those obtained from conventional econometric analysis in this literature. Results so obtained are highly relevant for efforts to promote inbound tourism by policy-makers as well as for efforts to stimulate increased tourist expenditures.

Eugenio-Martin and Inchausti-Sintes (2016) identify two key characteristics of low cost travel viz. longer stays and greater expenditure. They argue that by spending less on travel visitors are likely to spend more at their travel destination. Ferrer-Rosell and Coenders (2017) disagree with this argument, and their research concludes, as we do, that budget airline users tend to spend less. Furthermore, it can be argued that the difference between airline types in terms of expenditure by passengers is growing wider(Ferrer-Rosell and Coenders, 2017). At the same time, industry wide developments suggest that business models of both airline types are converging (Morlotti et al., 2017; Dobruszkes et al., 2017; Ferrer-Rosell and Coenders, 2017), as LCCs move from their old out-of-town regional airports to the main hubs[1] (Dobruszkes et al., 2017; Ferrer-Rosell and Coenders, 2017) and more airlines adopt dynamic pricing techniques pioneered by the LCCs. There is also an increasing trend toward yield maximisation by charging for services which were previously free such as prior seat allocations, checked in baggage and paid for meals.

Travellers using LCCs may also be expected to have other characteristics which predispose them to shorter stays. For example, younger age groups tend to spend less (Brida and Scuderi, 2013; Marrocu et al., 2015). Past research has modelled airline choice from observed flying patterns which does show some differences in passenger behaviour, but not to any major extent (Hess et al., 2007; Castillo-Manzano and Marchena-Gomez, 2010; Clavé et al., 2015). In addition, the income of the LCC passengers is often similar to those using other carriers(Hess et al., 2007). Many other choice factors influencing choice of air carrier have been considered, with a common example being in flight service as in (Han, 2013; Fourie and Lubbe, 2006; Cho et al., 2017). Our research makes use of nationality to control for income, an approach commonly adopted where individual income data is unavailable (Belenkiy and Riker, 2012; Eugenio-Martin and Campos-Soria, 2014) (Ali et al., 2016). Whilst the broad trend of the global economy has been upward through the period studied, the continued threat of a “double-dip ”recession means many passengers remain inclined toward LCCs (Bronner and de Hoog, 2014; Campos-Soria et al., 2015). Income also affects through satisfaction with the service level of LCCs, and hence the likelihood of wealthier travellers able to spend more using LCCs in ways which differ by nationality (Ali et al., 2016).

[1]Dobruszkes et al. (2017) provides the example of the change by Ryanair and Easyjet from Girona (GIR) to the main Barcelona (BCN) airport which is also used by the flag carriers.

Controlling for year deals with observed trends and this approach is used by us in line with the tourism literature surveyed by Brida and Scuderi (2013),Thrane (2014) and Dogru et al. (2017).

Two key dependent variables are of particular relevance: the length of time tourists stay in a destination and the amount of money which they spend in the local economy when they visit as tourists (Dogru et al., 2017). A number of robust methodologies have been proposed in the literature to address possible endogeneity issues (Thrane, 2015; Eugenio-Martin and Inchausti-Sintes, 2016). Use of per-day expenditure is also proposed as a solution in the literature (Sun and Stynes, 2006). It is conventional in this literature to focus mainly on total expenditure, while employing length of stay as an important explanatory variable (Brida and Scuderi, 2013; Thrane, 2014; Dogru et al., 2017). We employ stay duration as an explanatory variable for total expenditure, particularly since exogeneity is not a significant concern while making use of the counterfactual decomposition technique.

Within the literature on tourism expenditure there is an increasing acknowledgement of the importance of considering distributional effects rather than focusing on the mean alone (Almeida and Garrod, 2017; Marrocu et al., 2015). However within the length of stay literature focus on the mean tends to continue to dominate Thrane (2016a). Stay duration is measured in days and therefore not continuous like expenditure. Further the time that people stay has a tendency to focus on certain durations, for example on week or two weeks are far more common than 9 or 11 days say[2]. Because of these more common mass points distributions are referred to as having focal points. Logistic regression (Thrane, 2016b) and survival models (Barros et al., 2010; Wang et al., 2012; Ferrerrosell et al., 2014; Gémar et al., 2016) have also been utilised to recognise the focal point nature of the empirical distribution, but these do not provide effects across the distribution. Most travel is undertaken for short-stays, typically one week, two-weeks or one-month with very few stays having stay durations outside these. One of the major advantages of the Chernozhukov et al. (2013) approach is the ability to use a suitable distribution for a variable such as length of stay that is not continuous. Hence, whilst quantile regression (Koenker and Bassett Jr, 1978) and its subsequent development into unconditional quantile regression (Fortin et al., 2009), have continued relevance in continuous distributions, the approach we employ has wider validity. Consequently we are able to make a further contribution to the literature by assessing the distributional aspects of stay duration whilst simultaneously being a better measure of the treatment of LCC use.

This paper disentangles the effects of budget airline travel on expenditure and stay duration of visitors to the United Kingdom. Employing International Passenger Survey (UK ONS) data provides us with a number of key visitor characteristics. In combination with data on airline offers to passengers and the UK Consumer Price Index (CPI), we construct a dataset that enables the analysis of expenditure and length of stay for a five year period between 2011 and 2015 inclusive. We consider both the expenditure of tourists while in the UK and the duration of their stay. The latter assessment is made possible because our chosen methodology deals with dependent variables which have focal-point distributions as discussed. We focus on the European market as this offers one the largest number of available budget flights. We attempt to explain the observed differential in dependent variable distribution between budget airline travellers and those who arrive on full-service airlines using relevant variables, and we evaluate results across the full range of each distribution. We decompose this differential into that which is driven by differences in characteristics of the two samples (budget and full service airline passengers), and that which stems

[2]We demonstrate this with our data in Figure 1

structurally from airline type. By isolating the structural effect we provide a much clearer insight into the role of airline type on length of stay and expenditure allowing us to make a contribution to the literature.

The Chernozhukov et al. (2013) technique has primarily been employed in labour economics where wage distributions and inequality are big concerns (Depalo et al., 2015; Selezneva and Van Kerm, 2016). In addition to the decomposition of observed distributional differentials into structural and characteristic components, a number of "what if?" questions may be posed. We consider the effect of adding an additional nights stay, or the effect of proportionally extending a tourist's stay, on expenditure as an example of variable transformation. Given national differentials identified elsewhere in our analysis, decomposition enables us to demonstrate the impact of nationality more clearly. Our research thus offers three key contributions. We are the first to apply distributional regression decompositions in transportation and tourism, with the associated benefits outlined. Secondly this is the first study of the impact of LCCs on expenditure to consider distributional aspects from any methodology. Finally facilitated by the ability of Chernozhukov et al. (2013) to cope with focal point distributions this is the first study to consider length of stay away from average effects from any perspective. Consequently we are able to provide additional insights for policy analysis and illustrate the potential benefits of promoting inbound tourism from specific markets as well as encouraging increased lengths of stay.

The remainder of the paper is organised as follows. Section 2 introduces the IPS dataset that provides the empirical backdrop to the study, with Section 3 outlining the methodology that we employ. Results are disucssed in Section 4, followed by Section 5 which considers the implications of our main results. Section 6 concludes and it also highlights suggestions for promotion of inbound expenditure and stay duration, and the role of types of airlines.

# 2 Data

Our data comes from the United Kingdom International Passenger Survey (IPS) drawing on data released for the years 2011, 2012, 2013, 2014 and 2015 (Office for National Statistics, 2012, 2013, 2014, 2015, 2016). Because of the referendum on British membership of the European Union in June 2016 we do not include 2016 within our analysis. Similarly, data from earlier years, prior to 2011, witnessed stronger impact from the global financial crisis (2007-2009) and so those years are also omitted. We focus on passengers holding European nationality, either as members of the EU or in European countries outside the Schengen borderless region.

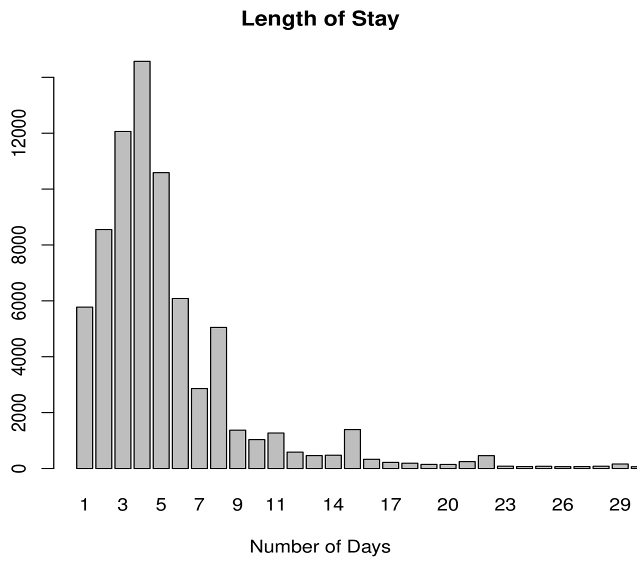
Table 1 provides details of the variables that are included representing various characteristics of surveyed passengers. In each case a $t$-test of mean equality is reported to identify differentials between LCC fliers and non-LCC fliers. It is clearly noticeable that passengers using LCCs show lower expenditures as well as longer stay durations. We find highly significant differences of almost £200 in average spending as well as one extra night being added to stay duration for such travellers. A larger proportion of LCC passengers are female and are under 25 years of age. Those using LCCS are highly likely to report purposes of travel mainly as going on holiday or visiting relatives, whilst business travellers make up a larger proportion of the other carriers' passenger markets. However, as previously noted, this situation is changing as business models are converging for LCCs and non-LCCs (Ferrer-Rosell and Coenders, 2017; Morlotti et al., 2017). 15.5% of budget travellers were on business in the 2011 to 2015 period. Whilst the vast majority of respondents

Table 1: Summary statistics for characteristics variables

| Variable | Mean | s.d. | Min | Max | Budget | Other | Difference |
|---|---|---|---|---|---|---|---|
| Expenditure (at 2015 prices) | 500.00 | 1109.21 | 0.996 | 150617 | 412.34 | 611.38 | -199.03 |
| Log expenditure (at 2015 prices) | 5.576 | 1.229 | -0.004 | 11.922 | 5.420 | 5.774 | $-0.355^{***}$ |
| Length of stay(days) | 6.796 | 13.493 | 1 | 355 | 7.260 | 6.206 | $1.054^{***}$ |
| Log length of stay (days) | 1.475 | 0.810 | 0.000 | 5.872 | 1.534 | 1.400 | $0.134^{***}$ |
| Male | 0.558 | 0.497 | 0.000 | 1.000 | 0.527 | 0.597 | $-0.070^{***}$ |
| Age (Proportions of total in each size) | | | | | | | |
| Aged 0-24 years | 0.171 | 0.376 | 0.000 | 1.000 | 0.214 | 0.116 | $0.098^{***}$ |
| Aged 25-64 years | 0.779 | 0.415 | 0.000 | 1.000 | 0.728 | 0.844 | $-0.116^{***}$ |
| Purpose of visit (Proportions of total in each size) | | | | | | | |
| Business | 0.251 | 0.433 | 0.000 | 1.000 | 0.155 | 0.372 | $-0.217^{***}$ |
| Holiday | 0.359 | 0.480 | 0.000 | 1.000 | 0.403 | 0.303 | $0.099^{***}$ |
| Visiting friends or relatives | 0.242 | 0.428 | 0.000 | 1.000 | 0.308 | 0.158 | $0.149^{***}$ |
| Other visitors (r) | 0.148 | 0.356 | 0.000 | 1.000 | 0.135 | 0.166 | $-0.032^{***}$ |
| Group Size (Proportions of total in each size) | | | | | | | |
| 1 (r) | 0.582 | 0.493 | 0.000 | 1.000 | 0.521 | 0.659 | $-0.138^{***}$ |
| 2 | 0.259 | 0.438 | 0.000 | 1.000 | 0.291 | 0.219 | $0.071^{***}$ |
| 3 | 0.070 | 0.256 | 0.000 | 1.000 | 0.081 | 0.057 | $0.024^{***}$ |
| 4 | 0.062 | 0.240 | 0.000 | 1.000 | 0.073 | 0.047 | $0.027^{***}$ |
| 5 or more | 0.027 | 0.163 | 0.000 | 1.000 | 0.034 | 0.018 | $0.017^{***}$ |
| Year (Proportions of total in each year) | | | | | | | |
| 2011 (r) | 0.191 | 0.393 | 0.000 | 1.000 | 0.192 | 0.189 | 0.003 |
| 2012 | 0.203 | 0.403 | 0.000 | 1.000 | 0.200 | 0.208 | $-0.008^{***}$ |
| 2013 | 0.209 | 0.407 | 0.000 | 1.000 | 0.203 | 0.217 | $-0.014^{***}$ |
| 2014 | 0.197 | 0.398 | 0.000 | 1.000 | 0.193 | 0.203 | $-0.010^{***}$ |
| 2015 | 0.200 | 0.400 | 0.000 | 1.000 | 0.212 | 0.184 | $0.028^{***}$ |
| Region of origin (Proportions of total in each region) | | | | | | | |
| European Union | 0.948 | 0.222 | 0.000 | 1.000 | 0.966 | 0.925 | $0.042^{***}$ |
| Eurppe: Non-European Union | 0.052 | 0.222 | 0.000 | 1.000 | 0.034 | 0.075 | $-0.042^{***}$ |

Notes: Summary statistics calculated for groups of variables such that means give proportions of total belonging to that class. (r) denotes reference category for each group. Significance reported for two-sample t.test of equality of means betwen LCC and non-LCC carriers. Source: Office for National Statistics (2012, 2013, 2014, 2015, 2016). Significance denoted by * - $p < 0.05$, **- $p < 0.01$ and *** - $p < 0.001$

Figure 1: Bar chart of stay duration for all respondents



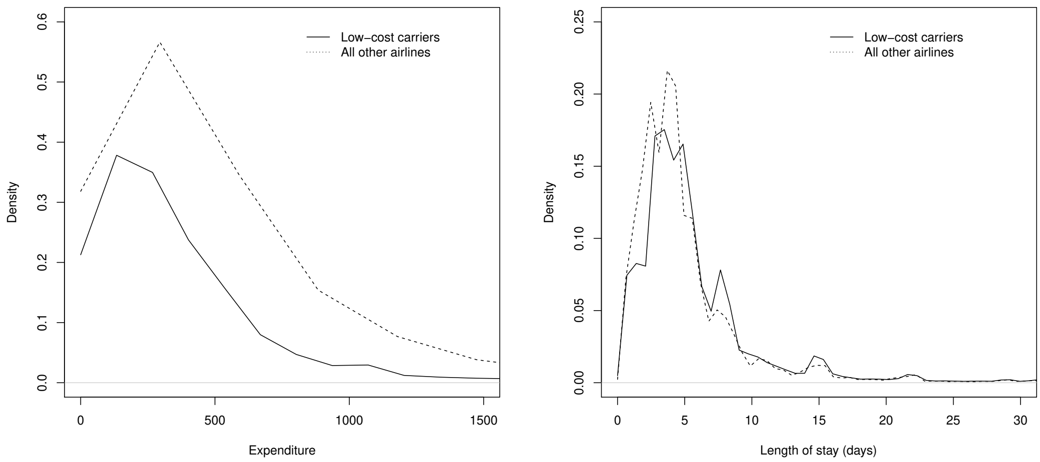

Notes: Bar chart omits stays longer than one month for clarity. Source: User calculations on Office for National Statistics (2012, 2013, 2014, 2015, 2016).

were travelling alone at every group size, the proportion for budget airlines is higher. Growth for LCCs is confirmed by the significantly higher proportion of total passenger numbers in 2015, compared to more even spreads for other carriers. Despite trends favouring growth within LCCs in 2015, overall figures do not appear to support the suggestion of Bronner and de Hoog (2014) and Campos-Soria et al. (2015) that airline passengers choose budget carriers through fear of future recessions. We also observe a bigger proportion of LCC users coming from the European Union.

To illustrate the focal point nature of the length of stay distribution Figure 1 provides a bar chart. A trip which begins on Saturday and then ends the following Saturday would cover seven nights and hence eight days. A clear spike at eight days is noted with almost twice as many respondents reporting staying eight days than seven, and nine day stays being just a quarter of the eight day figure. Likewise a two week stay covers fifteen days and a very notable increase in frequency is seen there. 22 and 29 days also stand out, again because they are whole numbers of weeks. Nevertheless the majority of stays are in the shorter number of days. Henceforth we refer to these spikes as focal points, recognising that they mark significant deviations from a smooth curve.

Figure 2 provides density plots of the two explanatory variables illustrating the differences between low-cost carrier fliers and those who use other airlines. Within the left panel there is clear evidence that low-cost carrier passengers spend less across the entire distribution. This contradicts the hypothesis proposed by Eugenio-Martin and Inchausti-Sintes (2016) that respondents substitute air fare for spending at their final destination. However, there is some evidence that budget airlines are successfully encouraging a slightly longer stay duration. Still clear on the diversity plots are the peaks at eight, fifteen and twenty-two days; these being the primary focal points referred to. This is also implicitly consistent with spending per day by budget airline travellers being lower, but this is not considered within our paper owing to its limited economic applicability. Both density plots

Figure 2: Impact of Budget Airline arrival on inbound tourist expenditure




Notes: Densities are plotted using all available data for each subsample, low-cost carriers ($n = 54,573$) and other airline fliers ($n = 42,955$). Expenditure is plotted for £0 to £1500 to focus on the majority of observations; 4705 respondents reported expenditure higher than this. Source: User calculations on Office for National Statistics (2012, 2013, 2014, 2015, 2016).

show changing differentials between LCC and non-LCC fliers across the distribution suggesting that it is useful to conduct a quantile analysis.

Given increasing interest in the role of nationality in influencing inbound tourist behaviour we employ the counterfactual technique (Chernozhukov et al., 2013) to extract the effect of tourists originating from the top five nations from the European continent by total passenger numbers. Table 2 provides details of the numbers and proportions of travellers from each country. We list total visitor numbers in descending order for the five year sample, which shows that Germany is the largest inbound tourist market for the UK, with the Republic of Ireland, France, Italy and the Netherlands completing the top five originating countries for inbound tourists into the UK. Large differentials exist between the bigger travelling nations and those who are the biggest users of LCCs. For example Spain and the Republic of Ireland have huge additional usage of LCCs whilst Scandinavian nations show the opposite result.

Growth in the LCC sector is well documented (Dobruszkes et al., 2017; Ferrer-Rosell and Coenders, 2017). A comparison of the proportion of respondents reporting travel by budget airline has increased to more than half towards the end of our sample period. Table 3 details the number of respondents grouped by year and airline type. It should be noted that although there has been a recent increase in proportional share of LCCs, the level was almost 50% in 2011 and hence more data would be needed to confirm trends. We find little support for the explanations provided by Bronner and de Hoog (2014) and Campos-Soria et al. (2015) about additional growth in LCC usage by passengers as a result of perceived adverse economic climate. It is more plausible to conclude that there is both a greater willingness to accept LCCs as alternatives to traditional full service carriers, and a continuing trend suggesting full service carriers adopting pricing strategies associated with LCCs.

Table 2: Number of inbound passengers by nationality

| Country | All Passengers | | LCC Airline | | Other Airline | | Difference |
|---|---|---|---|---|---|---|---|
| | No. | % Total | No. | % LCC Fliers | No. | % Others | |
| Germany | 14523 | 14.89% | 7464 | 13.68% | 7059 | 16.43% | 405 |
| Ireland | 11481 | 11.77% | 10350 | 18.97% | 1131 | 2.63% | 9219 |
| Italy | 8270 | 8.48% | 5144 | 9.43% | 3126 | 7.28% | 2018 |
| France | 7735 | 7.93% | 4247 | 7.78% | 3488 | 8.12% | 759 |
| Netherlands | 7316 | 7.50% | 3801 | 6.96% | 3515 | 8.18% | 286 |
| Spain | 7259 | 7.44% | 5221 | 9.57% | 2038 | 4.74% | 3183 |
| Sweden | 5840 | 5.99% | 1309 | 2.40% | 4531 | 10.55% | -3222 |
| Norway | 4963 | 5.09% | 1098 | 2.01% | 3865 | 9.00% | -2767 |
| Poland | 4771 | 4.89% | 4020 | 7.37% | 751 | 1.75% | 3269 |
| Denmark | 3557 | 3.65% | 1259 | 2.31% | 2298 | 5.35% | -1039 |
| Finland | 2163 | 2.22% | 237 | 0.43% | 1926 | 4.48% | -1689 |
| Switzerland | 1819 | 1.87% | 1144 | 2.10% | 675 | 1.57% | 469 |
| Portugal | 1771 | 1.82% | 768 | 1.41% | 1003 | 2.34% | -235 |
| Austria | 1682 | 1.72% | 502 | 0.92% | 1180 | 2.75% | -678 |
| Belgium | 1615 | 1.66% | 1139 | 2.09% | 476 | 1.11% | 663 |
| Romania | 1364 | 1.40% | 982 | 1.80% | 382 | 0.89% | 600 |
| Czech Republic | 1318 | 1.35% | 854 | 1.56% | 464 | 1.08% | 390 |
| Greece | 1235 | 1.27% | 850 | 1.56% | 385 | 0.90% | 465 |
| Russia | 1164 | 1.19% | 194 | 0.36% | 970 | 2.26% | -776 |
| Turkey | 1143 | 1.17% | 124 | 0.23% | 1019 | 2.37% | -895 |
| Hungary | 1122 | 1.15% | 808 | 1.48% | 314 | 0.73% | 494 |
| Lithuania | 803 | 0.82% | 740 | 1.36% | 63 | 0.15% | 677 |
| Bulgaria | 631 | 0.65% | 377 | 0.69% | 254 | 0.59% | 123 |
| Slovakia | 546 | 0.56% | 438 | 0.80% | 108 | 0.25% | 330 |
| Malta | 477 | 0.49% | 134 | 0.25% | 343 | 0.80% | -209 |
| Cyprus | 468 | 0.48% | 250 | 0.46% | 218 | 0.51% | 32 |
| Iceland | 445 | 0.46% | 105 | 0.19% | 340 | 0.79% | -235 |
| Latvia | 432 | 0.44% | 347 | 0.64% | 85 | 0.20% | 262 |
| Croatia | 255 | 0.26% | 92 | 0.17% | 163 | 0.38% | -71 |
| Estonia | 227 | 0.23% | 127 | 0.23% | 100 | 0.23% | 27 |
| Slovenia | 225 | 0.23% | 182 | 0.33% | 43 | 0.10% | 139 |
| Ukraine | 216 | 0.22% | 54 | 0.10% | 162 | 0.38% | -108 |
| Other EU | 178 | 0.18% | 71 | 0.13% | 107 | 0.25% | -36 |
| Other Non-EU | 484 | 0.50% | 118 | 0.22% | 366 | 0.85% | -248 |

Notes: Percentages are expressed as proportion of total respondents holding the given nationality. Difference expressed as number of budget airline passengers less the number of users of other airlines. Figures exclude those arriving by Sea. Source: User calculations on Office for National Statistics (2012, 2013, 2014, 2015, 2016)

Table 3: Number of respondents using budget airlines by year

| Year | LCC | Other Carriers | Total | Budget Share (%) |
|---|---|---|---|---|
| 2011 | 6718 | 6918 | 13636 | 49.27% |
| 2012 | 7041 | 7655 | 14696 | 47.91% |
| 2013 | 6491 | 6817 | 13308 | 48.78% |
| 2014 | 8754 | 8704 | 17458 | 50.14% |
| 2015 | 9493 | 7888 | 17381 | 54.62% |

# 3 Methods

In any application of the (Chernozhukov et al., 2013) technique the total population is divided into groups $k \in \mathscr{K}$, and of these the outcome variable $S$ is known for $j \in \mathscr{J} \subseteq \mathscr{K}$. For each population $k$ there is an observed set of covariates $X_k$ with dimension $d_k$, from which we can identify the covariate distribution $F_{X_k}$.

In our analysis we consider two variables of interest viz. expenditure $EX$ and length of stay $LS$, but for expositional simplicity we define $S = \{EX, LS\}$[3]. What follows could be applied for any continuous, or multiple valued, dependent variable. Define $F_{S\langle j|k\rangle}$ as the conditional distribution of $S$ when the sample $j$ has the covariate distribution of $k$. For population $j$ the conditional distribution of the outcome variable $S$ is given as $F_{S_j|X_j}$, where $X_j$ is the set of covariates as observed in sample $j$. We then write the distribution of the counterfactual outcome $F_{S\langle j|k\rangle}$, and its left inverse function as $F^{\leftarrow}_{S\langle j|k\rangle}$ allows combination of the conditional distribution of population $j$ with the covariate distributions of population $k$ to produce the counterfactual distribution as per (1) and quantile function (2).

$$F_{S\langle j|k\rangle} := \int_{\mathscr{X}_k} F_{S_j|X_j}(s|x)\, dF_{X_k}(x), \quad s \in \mathscr{S} \tag{1}$$

$$Q_{S\langle j|k\rangle} := F^{\leftarrow}_{S\langle j|k\rangle}(\tau), \quad \tau \in (0,1) \tag{2}$$

Equation (1) provides the distribution of the counterfactual outcome $S\langle j|k\rangle$ created by sampling the covariate $X_k$ from $F_{X_k}$ and then sampling the outcome $S\langle j|k\rangle$ from the conditional distribution $F_{S_j|X_j}(\cdot|X_k)$. $\mathscr{S}$ represents the set of possible alternative groups, here it is just 2 and $\mathscr{X}$ is the support of the distributions of $X_k$. Equation (3) provides the strong representation of this process and is useful for linking to regression models such as conditional quantile technique of Koenker and Bassett Jr (1978).

$$S\langle j|k\rangle = Q_{S_j|X_j}(U|X_k) \tag{3}$$

Where $U \sim U(0,1)$ independently of $X_k$ $F_{X_k}$.

Three types of counterfactual effects (CE) are defined by Chernozhukov et al. (2013). First the CE of changing the conditional distribution (4). Secondly, the effect of changing the covariate distribution whilst holding the conditional distribution constant, (5). Finally both may be combined

[3]This marks a change to the Chernozhukov et al. (2013) exposition in which $Y$ is used as the conventional representation of the dependent variable.

to produce equation (6).

$$\text{CE1:}\, Q_{S\langle j|k\rangle}(\tau) - Q_{S\langle l|k\rangle}(\tau) \tag{4}$$

$$\text{CE2:}\, Q_{S\langle j|k\rangle}(\tau) - Q_{S\langle j|m\rangle}(\tau) \tag{5}$$

$$\text{CE3:}\, Q_{S\langle j|k\rangle}(\tau) - Q_{S\langle l|m\rangle}(\tau) \tag{6}$$

Through this approach we can understand more about the impacts of the various covariates in $X_k$ by performing transformations upon them and utilising the CE2 effect process. In the main this approach permits the identification of counterfactual effects and enables the impact of a treatment to be decomposed into the true effect of treatment on the treated, and that observed difference attributable to the differences in characteristics between the treated and non-treated. In a simple case where $j = 1$ and $k = 0$, such that there are two groups, we can employ the counterfactual $F_{S\langle 0|1\rangle}$ to write equation (7).

$$F_{S\langle 1|1\rangle} - F_{S\langle 0|0\rangle} = \left[F_{S\langle 1|1\rangle} - F_{S\langle 0|1\rangle}\right] + \left[F_{S\langle 0|1\rangle} - F_{S\langle 0|0\rangle}\right] \tag{7}$$

Following Oaxaca (1973) and Blinder (1973) this expression gives the observed differential as the sum of the structural effect of being in population 1 and the characteristic differences between the two populations. These are then the structural and characteristic effects referred to in the discussion that follows.

An important note is necessary for both the exogeneity of the treatment policy and the exogeneity of the independent variables within $X_k$. In order to describe the structural effect as causal it is necessary to assume that individuals are randomly assigned to the treatment. In the case of covariate exogeneity it is required that changes to the $X_k$ variables, to $X_m$ for example, do not change the allocation process, that is $F_{X_m} = F_{X_k}$.

In our study of budget airlines both the length of stay and expenditure are known for all respondents making $\mathscr{J}$ and $\mathscr{K}$ equivalent. Indeed with only two populations to be considered, low-cost passengers compared to all other airlines, we have $\mathscr{J} = 2$. $j = 0$ represents the normal airlines and $j = 1$ denotes low cost carriers, aligning with the values of our dummy variable for budget airline use. Since passengers choose their flights it is unreasonable to consider our treatment as random and therefore the discussion of decompositions does not incorporate a causal interpretation. In contrast, in our examination of covariate changes it is reasonable to view the allocation process of characteristics to expenditure as being unaltered. For example the commonly observed[4] higher expenditure from individual travellers continues to apply.

Our estimations utilise the *R* package `counterfactual` (Chen et al., 2016) which is based upon the Chernozhukov et al. (2013) paper. In our results we utilise the quantile regression approach, invoking `quantreg` (Koenker, 2016), but for robustness we also exploit the location and scale shift using censored quantile regressions (Chernozhukov and Hong, 2002), [5]. Estimates under the quantile specification are given by:

$$\hat{F}_{S_j,X_j}(s|x) = \varepsilon + \int_{\varepsilon} 1\{x'\hat{\beta}_j(u) \leq s\} \tag{8}$$

[4]This appears in Brida and Scuderi (2013), Thrane (2014), Marrocu et al. (2015) and Almeida and Garrod (2017) amongst many others.

[5]Descriptions of all of these are available in the accompanying documentation to (Chen et al., 2016).

in which $\hat{\beta}_j(u)$ is the Koenker and Bassett Jr (1978) estimator:

$$\hat{\beta}_j(u) = \arg\min_{b\in\mathbb{R}^{d^x}} \sum_{i=1}^{n_j} \left[u - 1\{S_{ji} \le X'_{ji}b\}\right] \left[S_{ji} - X'_{ji}b\right]$$

This method employs a trimming parameter $\varepsilon$ to estimate its coefficients and is suitable only for continuous $S$.

In our empirical study expenditure is continuous whilst the length of stay necessarily is measured in days limiting the set of values slightly. Hence we adopt the logistic distribution for stay duration, as follows:

$$\hat{F}_{S_j|X_j}(s|x) = \Lambda\left(x'\hat{\beta}(s)s\right) \tag{9}$$

$$\hat{\beta}(s) = arg\min_{b\in\mathbb{R}^{d^x}} \sum_{j=1}^{n_j} \left[1\{S_{ji} \le s\} log\Lambda\left(x'_{ij}b\right) + 1\{S_{ji} \ge s\} log\Lambda\left(x'_{ij}b\right)\right] \tag{10}$$

In (9) and (10) $S_{ji}$ is the observation on dependent variable $S$ for individual $i$ on variable $j$. Chen et al. (2016) note that this estimator has the flexibility to provide different parameters at different parts of the distribution. Against this specification robustness is tested using the probit and linear probability models.

# 4 Results

Through application of the counterfactual distribution methodologies we seek to address three key questions. First, we decompose the observed differentials between LCC passengers and other airline passengers on both length of stay and expenditure, assessing the extent to which previously implied lower spending and longer stay durations might be structurally attributed to airline type. Secondly, we explore the role of nationality in greater depth and we show what happens if the nations who provide the greatest number of inbound tourists to the UK were to behave at the European Union average. Finally we consider the length of stay, presenting results arising when stay is extended by a day, or proportionally to observed duration.

## 4.1 Decomposition of LCC Differentials

LCC passengers stay longer but spend less than their peers travelling by non-LCC carriers, but these results are drawn from the mean of the distribution (Eugenio-Martin and Inchausti-Sintes, 2016; Ferrer-Rosell and Coenders, 2017). While estimating our models, we are able to test for constancy of impact on expenditure and stay duration by using Kolmogorov-Smirnov (KS) and Cramer-von-Misses-Smirnov (CMS) tests. Results from our decomposition are plotted graphically in Figure 3. Test results are reported in Table 4.

Table 4 reports first the test for the significance of the overall, structural and characteristic effects in both the expenditure and length of stay studies. In all cases the probability of zero effect is almost nil. Having established a non-zero effect the second test is of the null hypothesis of constant effect across the distribution, that $QE(\tau) = QE(0.5) \forall \tau \in (0.1, 0.9)$, rejecting in all but the CMS

Table 4: Diagnostic tests for counterfactual inference

| Test | Length of Stay | | | | | | Expenditure | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total Effect | | Structural Effect | | Characteristic Effect | | Total Effect | | Structural Effect | | Characteristic Effect | |
| | KS | CMS | KS | CMS | KS | CMS | KS | CMS | KS | CMS | KS | CMS |
| No Effect: $QE(\tau) = 0 \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Constant Effect: $QE(\tau) = QE(0.5) \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.09 |
| Stochastic Dominance: $QE(\tau) > 0 \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0.65 | 0.85 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stochastic Dominance: $QE(\tau) < 0 \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0.81 | 0.81 | 0.87 | 0.87 | 0.49 | 0.65 |

Notes: Figures represent p-values for the tests of no effect, constant effect and stochastic dominance. KS is the Kolmogorov-Smirnov test, whilst CMS denotes the Cramer-von-Misses-Smirnov test. All statstics are generated using Chen et al. (2016). Full details of the tests may be found in Chernozhukov et al. (2013).

test for the composition effect in the length of stay regression. Policy-makers and practitioners alike are interested in whether the difference and effects increase, or decrease, expenditure and stay duration. The stochastic dominance tests evaluate this. In all cases expenditure is negative meaning that the null hypothesis of $QE(\tau) < 0 \forall \tau \in (0.1, 0.9)$ is not rejected. For the length of stay Figure 3 displays positive and negative significance at different $\tau$ meaning both $QE(\tau) < 0 \forall \tau \in (0.1, 0.9)$ and $QE(\tau) > 0 \forall \tau \in (0.1, 0.9)$ are rejected. Only the composition effect for stay duration is seen to be significantly positive, accepting $QE(\tau) > 0 \forall \tau \in (0.1, 0.9)$.

Average effects do not take into account characteristics of the fliers who use the LCCs. Figure 3 contains two plots, the upper three panels showing (a) overall, (b) structural and (c) composition effects for the stay duration and the lower panel, (d)-(f), showing the same for log expenditure. For stay duration the total effect has small negative region around the $\tau = 0.25$ and $\tau = 0.6$ levels then is positive from $\tau = 0.65$ onwards. In the central plot the structural effect is shown to also display regions of significant negative response, but these are much larger and more frequent. Above $\tau = 0.8$ the structural effect is negative, opposing the overall effect and leaving the composition effect to explain the observed longer stay duration. This implies that the overall effect is masking stay duration reductions across the distribution, and producing incorrect policy conclusions about the ability of LCC use to promote longer stay durations. Following from this analysis, it is still true that budget airline travellers stay longer as shown in Figure 2, but often this is due to differences in the observed samples rather than due to the use of an LCC. Table 4 confirms the significance of the above result and shows that there are no constant effects across the quantiles. Given the demonstrated variation across quantile both Figure 3 and Table 4 confirm that applying distributional regression analysis for stay duration is useful for both drawing additional insights and better policy formulation [6]. The application of Chernozhukov et al. (2013) enables this.

Figure 2 demonstrated that for all points on the distribution, LCC passengers spend less within the United Kingdom. When decomposing this it is the structural effect that dominates, especially at the higher end of the spending distribution. This dominance informs that all else equal LCC fliers can be expected to spend less; there is something within the choice to use an LCC that later manifests itself as lower spending but is not picked up by any of our controls. From all other characteristics captured by the dataset the right hand plot identifies only a slightly lower expenditure would be predicted. Whilst nationality controls for income partially it does not do so fully and it thus represents a characteristic that is likely to reduce the structural effect magnitude. In this case the decomposition reaffirms the role of budget airlines as reducing expenditure that regression analyses by **?** and others had picked up. However, we also show that although the sign is always negative the magnitude varies across the distribution, doing so for the first time. Being able to identify the structural effect is thus informative about the likely role of LCC use within expenditure and stay duration, and help provide additional insights within the area of research.

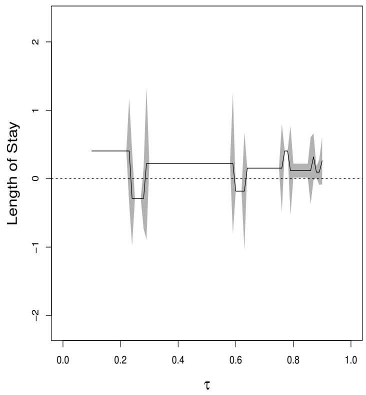
## 4.2 National Effects

Previous works reviewed in Brida and Scuderi (2013) and Dogru et al. (2017) have identified clear differentials in the way that different nationalities treat expenditure. This is something which can be readily identified by employing the counterfactual technique to subsamples within the European

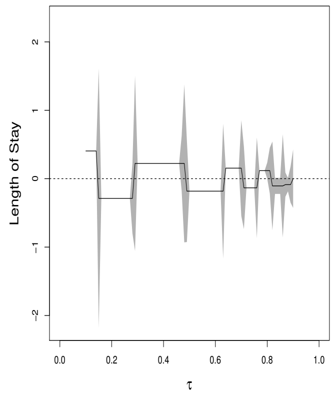
[6]In the case of expenditure the characteristic effect the CMS test rejects the null hypothesis at 95% significance but accepts at 90%.

Figure 3: Impact of budget airline arrival on inbound tourist behaviour

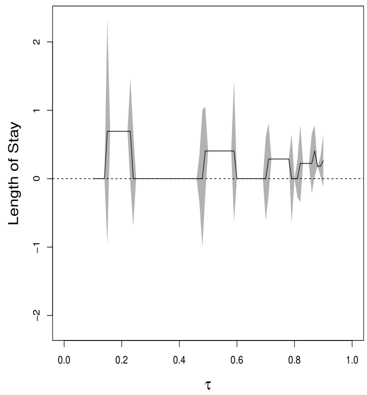
(a) Stay: Total



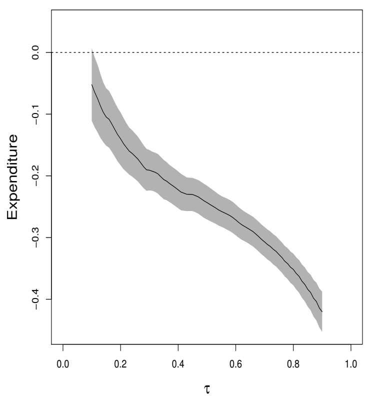

(b) Stay: Structural



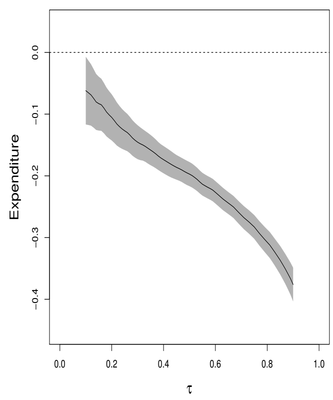

(c) Stay: Characteristics



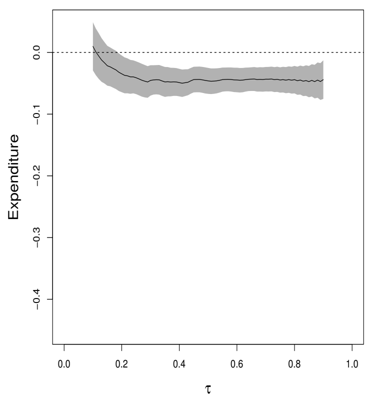

(d) Expenditure: Total




(e) Expenditure: Structural




(f) Expenditure: Characteristics




Distribution estimates plotted as solid lines with 95% confidence intervals as grey polygons. Total effect, $\Delta_O$, in left panel, structural effect ($\Delta_S$) attributed to budget airlines in centre panel and composition effect ($\Delta_X$) in the right hand panel. All estimates using Chen et al. (2016)..

Table 5: Nationality counterfactual impacts on log expenditure

| Nationality | Quantiles $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ | $\tau = 0.75$ | $\tau = 0.9$ |
|---|---|---|---|---|---|
| Germany | 0.036 | 0.043 | 0.053 | 0.076 | 0.104 |
| | (0.008) | (0.006) | (0.005) | (0.006) | (0.008) |
| Ireland | 0.042 | 0.045 | 0.043 | 0.046 | 0.054 |
| | (0.008) | (0.005) | (0.004) | (0.004) | (0.003) |
| Italy | 0.043 | 0.038 | 0.035 | 0.039 | 0.050 |
| | (0.005) | (0.003) | (0.003) | (0.002) | (0.004) |
| France | 0.005 | 0.012 | 0.021 | 0.039 | 0.059 |
| | (0.004) | (0.003) | (0.003) | (0.004) | (0.005) |
| Netherlands | 0.032 | 0.032 | 0.030 | 0.035 | 0.043 |
| | (0.005) | (0.003) | (0.002) | (0.003) | (0.004) |

Notes: Effects measured as type CE2 given in equation (5). Estimations using Chen et al. (2016).

Table 6: Diagnostic tests for counterfactual inference

| Test | Germany KS | Germany CMS | Ireland KS | Ireland CMS | Italy KS | Italy CMS | France KS | France CMS | Netherlands KS | Netherlands CMS |
|---|---|---|---|---|---|---|---|---|---|---|
| No Effect: $QE(\tau) = 0 \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Constant Effect: $QE(\tau) = QE(0.5) \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0.04 | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stochastic Dominance: $QE(\tau) > 0 \forall \tau \in (0.1, 0.9)$ | 0.77 | 0.77 | 0.81 | 0.81 | 0.80 | 0.80 | 0.83 | 0.83 | 0 | 0 |
| Stochastic Dominance: $QE(\tau) < 0 \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

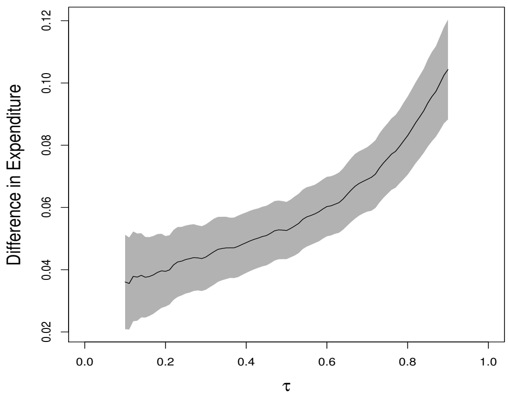
Notes: Figures represent p-values for the tests of no effect, constant effect and stochastic dominance. KS is the Kolmogorov-Smirnov test, whilst CMS denotes the Cramer-von-Misses-Smirnov test. All statstics are generated using Chen et al. (2016). Full details of the tests may be found in Chernozhukov et al. (2013).

population. For this exercise we consider only the largest five countries in terms of visitors between 2011 and 2015, viz. Germany, Republic of Ireland, Italy, France and the Netherlands. All these countries are members of the European Union and the Euro currency area. We plot the impacts in Figure 4 and then report the tests of effect, parameter equality and stochastic dominance in Table 6

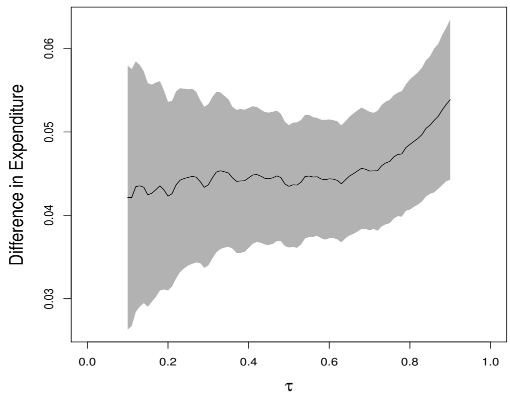
In each case the nationality variable produces an expenditure premium over other EU countries. This is perhaps unsurprising given the link between tourism and income. However, we also note some differences between the magnitude of the effect, particularly for the case of France, which shows much lower impacts at lower quantiles. Table 5 and Figure 4 both show that for lower $\tau$ values the effect for France is insignificant. All other values are highly significant confirming the importance of nationality, which effectively proxies tourist income. By looking at the tests in Table 6 we can see that all of the countries have stochastic dominance of $\Delta > 0$, including France in spite of the insignificance around the 10th percentile. Consequently in all five cases expenditure could be expected to be lower if the respondents did not have the unobserved characteristics of their particular nationality. All else equal Germans increase log expenditure by between 0.04 and 0.1, above the reference category of smaller EU nations. From a promotion perspective these results suggest targeting French nationals may be ineffective at increasing expenditure, raising only the upper end of the distribution. Irish, Dutch and Italian nationals deliver a similar premium across the distribution making blanket promotions to these nationalities equally effective irrespective of

Figure 4: Nationality effect on log tourism expenditure

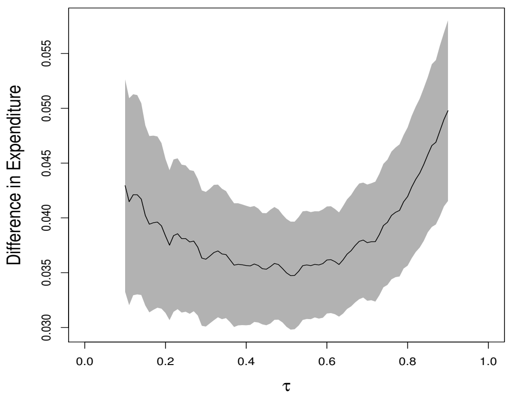
(a) Germany



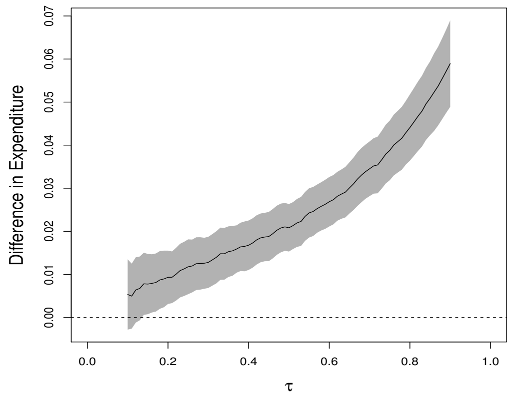

(b) Ireland



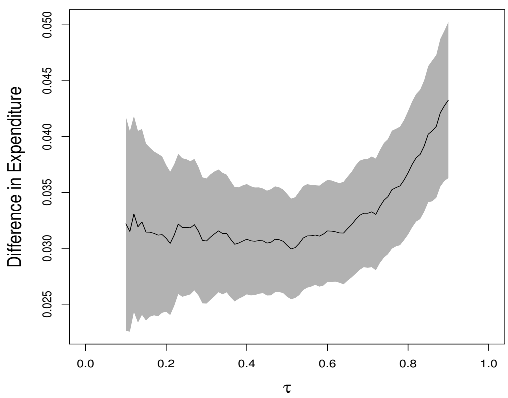

(c) Italy




(d) France



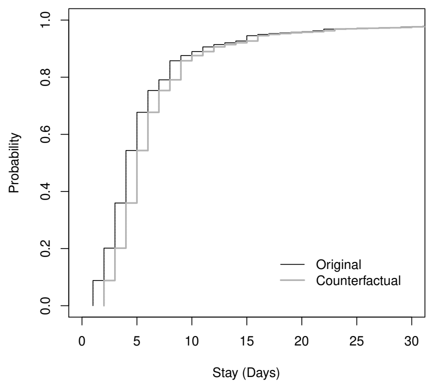

(e) Netherlands



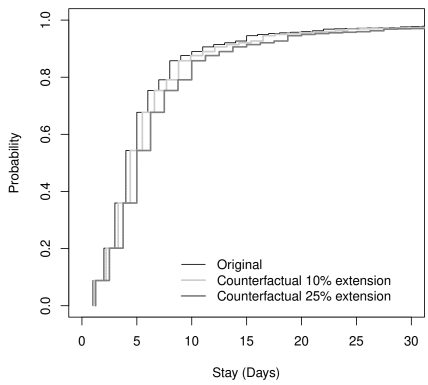

Distribution estimates plotted as solid lines with 95% confidence intervals as grey polygons. Total effect, $\Delta_O$, in left panel, structural effect ($\Delta_S$) attributed to budget airlines in centre panel and composition effect ($\Delta_X$) in the right hand panel. All estimates using Chen et al. (2016)..

Figure 5: Conditional density function with counterfactual distributions

(a) Extra day




(b) Proportional increase




their other characteristics. Overall the stochastic dominance of all five nations points to them all being sensible markets for advertising over and above other EU nations.

## 4.3 Extending Stay

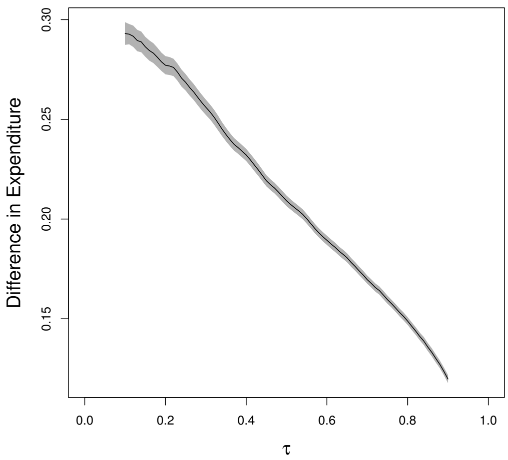
Given particular characteristics, allocation of spending is not expected to change if the length of stay is extended slightly. Hence, we consider what impact would be felt if the $X_k$ matrix is updated to $X_m$ with stay extended by one (i.e. one additional day's stay), and the logarithms are recalculated accordingly. As an additional day would have a larger proportional impact on short-stayers, we also consider two proportionate increases in spending, by 10% and 25% respectively, which are proportional to the amounts spent by tourists. Whilst the latter lacks intuition in terms of length of stay because time may be subdivided, it is not unreasonable to consider that respondents stay for fractions of days. We also propose extending the stay of those who stay for more than the typical long weekend, i.e. four days or more.

Figure 5 includes the plots of the actual number of days stayed for easier interpretation. Panel (a) shows the additional day stay case and panel (b) shows the proportionate increases. Within this figure the stepped nature of the length of stay is clearly identified and the dominance of short trips is clear. By comparing the two panels of Figure 5 the effect of a percentage increase versus a blanket increase is clear, the former making very little difference to the vast majority of respondents whose stay is less than four days. In the fourth scenario an additional day is added for respondents who stay four days or longer producing a counterfactual distribution which mirrors the real population for short stays then mimics the one extra day case from four days onwards. Consequently plotting this in Figure 5 would not indistinguishable from the other lines on the plot and is omitted from panel (a). Next we estimate the effects of these distribution changes plotting the results in Figure 6. We report the effects of each policy in Table 7 and the relevant tests in Table 8.

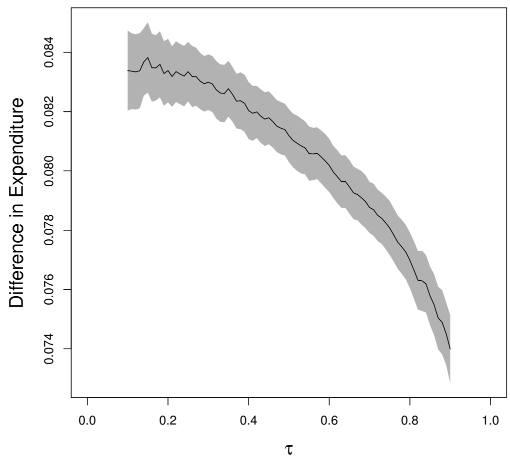
When all respondents are able to stay an extra day, this represents a bigger proportion of the time stated in the UK for the shortest stayers. Consequently a much larger magnitude of impact is noted at the lowest end of the total spending distribution. Panel (a) of Figure 6 informs us that at

Figure 6: Length of Stay Counterfactuals

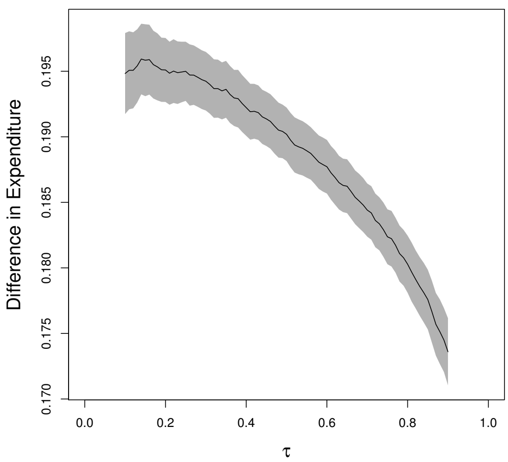
(a) Additional Day



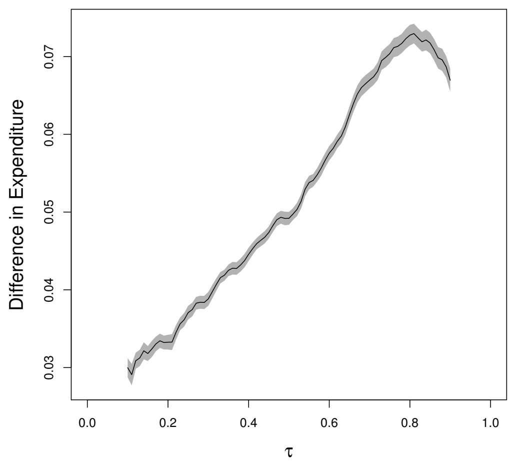

(b) Proportional increase: 10%




(c) Proportional increase: 25%




(d) Additional day on longer stays




Effect estimates plotted as solid lines with 95% confidence intervals as grey polygons. Total effect, $\Delta_O$, in left panel, structural effect ($\Delta_S$) attributed to budget airlines in centre panel and composition effect ($\Delta_X$) in the right hand panel. All estimates using Chen et al. (2016)..

Table 7: Stay extension counterfactual impacts on log expenditure

| Policy | Quantiles | | | | |
|---|---|---|---|---|---|
| | $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ | $\tau = 0.75$ | $\tau = 0.9$ |
| Extend by 1 day | 0.296 | 0.269 | 0.209 | 0.160 | 0.112 |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) |
| Extend by 10% | 0.083 | 0.083 | 0.081 | 0.078 | 0.074 |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.001) |
| Extend by 25% | 0.195 | 0.195 | 0.190 | 0.182 | 0.174 |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| Extend long stay by 1 day | 0.030 | 0.037 | 0.049 | 0.070 | 0.067 |
| | (0.001) | (0.000) | (0.000) | (0.001) | (0.001) |

Notes: Effects measured as type CE2 given in equation (5). Estimations using Chen et al. (2016).

Table 8: Diagnostic tests for counterfactual inference

| Test | Extend: 1 day | | Extend: 10% | | Extend: 25% | | Extend: long | |
|---|---|---|---|---|---|---|---|---|
| | KS | CMS | KS | CMS | KS | CMS | KS | CMS |
| No Effect: $QE(\tau) = 0 \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Constant Effect: $QE(\tau) = QE(0.5) \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stochastic Dominance: $QE(\tau) > 0 \forall \tau \in (0.1, 0.9)$ | 1 | 1 | 0.87 | 0.67 | 0.84 | 0.84 | 1 | 1 |
| Stochastic Dominance: $QE(\tau) < 0 \forall \tau \in (0.1, 0.9)$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Notes: Figures represent p-values for the tests of no effect, constant effect and stochastic dominance. KS is the Kolmogorov-Smirnov test, whilst CMS denotes the Cramer-von-Misses-Smirnov test. All statstics are generated using Chen et al. (2016). Full details of the tests may be found in Chernozhukov et al. (2013).

$\tau = 0.1$ an increase in log expenditure of almost 0.3 can be expected. However, once the increase is made proportional, this strong negative relationship weakens. Panels (b) and (c) reflect negative associations, but the strength of these is much lower. The bigger impact continues to fall at the lowest end of the expenditure distribution. Only in the case of an increase of stay by one day for longer stayers is the higher end of the distribution significantly increased. This appears to be intuitive given that this increase continues to represent a much bigger proportion for those staying less than a week than it does for the longest stayers.

Utilising these results to determine policy necessitates an understanding of which end of the distribution the government seeks to increase expenditure at (lower or upper). A blanket additional day, as well as proportional increases, generate more expenditures from the lowest spenders. An additional day for those staying more than four days impacts the most for high spenders. Closer inspection confirms that the area underneath the curve in panel (d) is the largest, such that the greatest total increase comes from this policy. However, these analyses are conditional on the *ceteris paribus* condition. In reality, it is unlikely that all respondents would maintain their per-day spending in the situations outlined above.

# 5 Discussion

Chernozhukov et al. (2013) counterfactual decomposition technique permits the identification of structural effects of variables across the distribution of interest by exploiting the hybrid case of the characteristics of one group being mapped to the variable of interest via the assignment function of another group. We demonstrated that the observed longer stays of LCC fliers are often driven by the characteristics of those passengers surveyed rather than the fact that they used an LCC. Indeed the structural effect of budget airlines were often to reduce stays, particularly towards the upper end of the stay distribution. We also demonstrate that lower spending within the UK attributable to passengers arriving on budget airlines is partially driven by the choice to take a cheaper flight. The controls employed in our analysis are those which are widely adopted in the tourism literature. These reflect the shortage of information about precise levels of income through nationality effects. We address this issues by carrying out estimation focusing on citizens of European countries alone. We show that the characteristics of LCC users lead to longer stays and lower spending, corresponding to a lower overall spend per day.

Budget airline use is mainly associated with short trips, whereby lower fares enable more weekend or city-break trips rather than longer but less frequent travel (Castillo-Manzano and Marchena-Gomez, 2010). Our methodology identifies this clearly within the negative regions of the structural effect, indeed the $\tau$ range of the negative structural effect is almost as large as that over which the structural effect is positive. Consequently the observed positive effect of LCC use, and that reported in the extant literature (Eugenio-Martin and Inchausti-Sintes, 2016; Ferrer-Rosell and Coenders, 2017), is attributable to the characteristics of the passengers rather than the type of airline used as claimed. The implication of this result is that encouraging more LCC use to generate longer stays is unlikely to be as effective as past research has suggested. By considering a number of counterfactual cases we assess impacts across the distributional of relevant variables to reveal the expected impact of policy changes; extending stay duration by one day being an example studied here. Increasing the duration of stay, unsurprisingly, does increase expenditure but we demonstrate that the additional expenditure primarily arises from lower overall spenders. This is partially driven by the higher spending per day of short stayers. This result remains unchanged when the extension is proportional to the stay duration. Generation of the greatest amount of additional spending relies on adding more time to the stay duration of shorter visitors.

Targeting promotion activities at particular nations has clear benefit. In our second counterfactual experiment we allow changes to the nationality characteristic to identify the true effect on expenditure attributable to nationality (and implicitly incomes from various nations). In each case, the largest providers of tourists to the United Kingdom showed greater amounts of spending than the smaller EU nations, but this happens differentially across the distribution. Our results clearly demonstrate that French inbound tourists do not spend as much as some others at the lower end of the distribution. Thus France is clear case where mainly higher total spenders should be targeted for tourism promotion efforts. By contrast, our other nations all display large increases to expenditure at the 10th percentile. Germans have an upward sloping effect such that the largest potential is amongst higher spenders, but unlike the French case, the magnitude of the effect is much greater. Ireland and, to a lesser extent Italy, have similar impacts across the range of $\tau$.

This technique has the potential to provide key distributional insights in cases where there are two obvious samples to compare such as LCC and non-LCC carriers in our case. In our research we refrain from discussing causality relating to choice of travel with LCCs. Inclusion in our LCC sample is driven by the characteristics of the respondents (Hess et al., 2007; Mason and Alamdari, 2007; Cho et al., 2017). However a similar approach can be applied in contexts where a clear choice has not been made, for example comparing genders or nationalities[7]. As we have distributional measures we can also consider issues of inequality. A commonly explored area for such analysis is labour economics where this method is most commonly applied and where the example within Chernozhukov et al. (2013) is set.

# 6 Conclusions

Budget airlines (low cost carriers) are an ever expanding part of the international transportation landscape and yet little systematic research has been carried out to understand their economic impact on key tourism variables, especially within a distributional context. Methodologically, length of stay has presented a challenge for empirical modelling as it is both discrete and has obvious focal points within its density function. We demonstrated the latter issue here in highlighting peaks in the density at eight, fifteen and twenty-two days providing focal points to those small neighbourhoods of the distribution. The discrete nature of the stay duration is immediate from the inability to have a non-integer number of nights stay. These features then render commonly applied distributional techniques, such as quantile regression, inadequate (Jones et al., 2015). Utilising the counterfactual decomposition method of Chernozhukov et al. (2013) we therefore robustly address an important gap in current knowledge whilst simultaneously resolving the issues of properly understanding distributional aspects of stay duration. Our results demonstrate that low cost carrier passengers spend less, particularly higher up the total expenditure range, and that they stay longer than those using other non-budget airlines. We address potential endogeneity of outcomes by removing the causal interpretation and focus on inferences which are methodologically robust and justifiable. In both cases studied (stay duration and nationality), significant variation in the coefficients is identified

[7] We would not be able to do the same with residence, since individuals do have some choice over where to live.

and the need to consider effects away from the mean confirmed.

We use data from the United Kingdom International Passenger Survey which provides a rich source of useful information. However, it has an important limitation in that respondents' income is not reported. We therefore employ nationality as a proxy for income. Attitudes towards LCCs and the choice to fly with them are closely linked to characteristics and cultural settings of survey respondents reinforcing nation as an appropriate level for control. We cannot conclude that a richer nation produces travellers with higher incomes, though we do find that the largest number of visitors come from the wealthiest nations. Our results show that expenditure arising from passengers using LCCs is lower than that of other non-LCC airlines. Our analysis does not model the choice to fly in the first instance owing to data limitations. We cannot infer that greater overall spending on flights would result if there were no LCC flights because many may people may simply choose not to travel in such a situation. Notwithstanding these concerns the results we present confirm that budget airlines provide significant value to an economy in their promotion of longer stays, higher expenditures and through generation of further inbound visitor numbers.

# References


Ali, F., Kim, W. G., and Ryu, K. (2016). The effect of physical environment on passenger delight and satisfaction: Moderating effect of national identity. *Tourism Management*, 57:213–224.

Almeida, A. and Garrod, B. (2017). Insights from analysing tourist expenditure using quantile regression. *Tourism Economics*, 23(5):1138–1145.

Barros, C. P., Butler, R., and Correia, A. (2010). The length of stay of golf tourism: A survival analysis. *Tourism Management*, 31(1):13–21.

Belenkiy, M. and Riker, D. (2012). Face-to-face exports the role of business travel in trade promotion. *Journal of Travel Research*, 51(5):632–639.

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4):436–455.

Brida, J. G. and Scuderi, R. (2013). Determinants of tourist expenditure: a review of microeconometric models. *Tourism Management Perspectives*, 6:28–40.

Bronner, F. and de Hoog, R. (2014). Vacationers and the economic double dip in europe. *Tourism management*, 40:330–337.

Campos-Soria, J. A., Inchausti-Sintes, F., and Eugenio-Martin, J. L. (2015). Understanding tourists' economizing strategies during the global economic crisis. *Tourism Management*, 48:164–173.

Castillo-Manzano, J. I. and Marchena-Gomez, M. (2010). Analysis of determinants of airline choice: profiling the lcc passenger. *Applied Economics Letters*, 18(1):49–53.

Chen, M., Chernozhukov, V., Fernandez-Val, I., and Melly, B. (2016). *Counterfactual: Estimation and Inference Methods for Counterfactual Analysis*. R package version 1.0.

Chernozhukov, V., Fernández-Val, I., and Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268.

Chernozhukov, V. and Hong, H. (2002). Three-step censored quantile regressions and extramarital affairs. *Journal of the American Statistical Association*, 97:872–882.

Cho, W., Windle, R. J., and Dresner, M. E. (2017). The impact of operational exposure and value-of-time on customer choice: Evidence from the airline industry. *Transportation Research Part A: Policy and Practice*.

Clavé, S. A., Saladié, Ò., Cortés-Jiménez, I., Young, A. F., and Young, R. (2015). How different are tourists who decide to travel to a mature destination because of the existence of a low-cost carrier route? *Journal of Air Transport Management*, 42:213–218.

Depalo, D., Giordano, R., and Papapetrou, E. (2015). Public–private wage differentials in euro-area countries: evidence from quantile decomposition analysis. *Empirical Economics*, 49(3):985–1015.

Dobruszkes, F., Givoni, M., and Vowles, T. (2017). Hello major airports, goodbye regional airports? recent changes in european and us low-cost airline airport choice. *Journal of Air Transport Management*, 59:50–62.

Dogru, T., Sirakaya-Turk, E., and Crouch, G. I. (2017). Remodeling international tourism demand: Old theory and new evidence. *Tourism Management*, 60:47–55.

Eugenio-Martin, J. L. and Campos-Soria, J. A. (2014). Economic crisis and tourism expenditure cutback decision. *Annals of Tourism Research*, 44:53–73.

Eugenio-Martin, J. L. and Inchausti-Sintes, F. (2016). Low-cost travel and tourism expenditures. *Annals of Tourism Research*, 57:140–159.

Ferrer-Rosell, B. and Coenders, G. (2017). Airline type and tourist expenditure: Are full service and low cost carriers converging or diverging? *Journal of Air Transport Management*, 63:119–125.

Ferrerrosell, B., Martnezgarcia, E., and Coenders, G. (2014). Package and no-frills air carriers as moderators of length of stay. *Tourism Management*, 42(42):114–122.

Fortin, N., Lemieux, T., and Firpo, S. (2009). Unconditional quantile regression. *Econometrica*, 77(3):953–973.

Fourie, C. and Lubbe, B. (2006). Determinants of selection of full-service airlines and low-cost carriersa note on business travellers in south africa. *Journal of Air Transport Management*, 12(2):98–102.

Gémar, G., Moniche, L., and Morales, A. J. (2016). Survival analysis of the spanish hotel industry. *Tourism Management*, 54:428–438.

Han, H. (2013). Effects of in-flight ambience and space/function on air travelers' decision to select a low-cost airline. *Tourism management*, 37:125–135.

Hess, S., Adler, T., and Polak, J. W. (2007). Modelling airport and airline choice behaviour with the use of stated preference survey data. *Transportation Research Part E: Logistics and Transportation Review*, 43(3):221–233.

Jones, A. M., Lomas, J., and Rice, N. (2015). Healthcare cost regressions: going beyond the mean to estimate the full distribution. *Health economics*, 24(9):1192–1212.

Koenker, R. (2016). *quantreg: Quantile Regression*. R package version 5.24.

Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, 46:33–50.

Marrocu, E., Paci, R., and Zara, A. (2015). Micro-economic determinants of tourist expenditure: A quantile regression approach. *Tourism Management*, 50:13–30.

Mason, K. J. and Alamdari, F. (2007). Eu network carriers, low cost carriers and consumer behaviour: A delphi study of future trends. *Journal of Air Transport Management*, 13(5):299–310.

Morlotti, C., Cattaneo, M., Malighetti, P., and Redondi, R. (2017). Multi-dimensional price elasticity for leisure and business destinations in the low-cost air transport market: Evidence from easyjet. *Tourism Management*, 61:23–34.

Oaxaca, R. (1973). Male-Female wage differentials in urban labour markets. *International Economic Review*, 14(3):693–709.

Office for National Statistics (2012). *International Passenger Survey 2011*. [data collection]. 6th Edition. UK Data Service. SN: 6846. DOI: http://doi.org/10.5255/UKDA-SN-6846-6.

Office for National Statistics (2013). *International Passenger Survey 2012*. [data collection]. 5th Edition. UK Data Service. SN: 7087 DOI: http://doi.org/10.5255/UKDA-SN-7087-5.

Office for National Statistics (2014). *International Passenger Survey 2013*. [data collection]. 3rd Edition. UK Data Service. SN: 7380. DOI: http://doi.org/10.5255/UKDA-SN-7380-3.

Office for National Statistics (2015). *International Passenger Survey 2014*. [data collection]. 4th Edition. UK Data Service. SN: 7534. DOI: http://doi.org/10.5255/UKDA-SN-7534-4.

Office for National Statistics (2016). *International Passenger Survey 2015*. [data collection]. 4th Edition. UK Data Service. SN: 7754 DOI: http://doi.org/10.5255/UKDA-SN-7754-4.

Selezneva, E. and Van Kerm, P. (2016). A distribution-sensitive examination of the gender wage gap in germany. *Journal of Economic Inequality*, 14(1):21.

Sun, Y.-Y. and Stynes, D. J. (2006). A note on estimating visitor spending on a per-day/night basis. *Tourism Management*, 27(4):721–725.

Thrane, C. (2014). Modelling micro-level tourism expenditure: recommendations on the choice of independent variables, functional form and estimation technique. *Tourism Economics*, 20(1):51–60.

Thrane, C. (2015). On the relationship between length of stay and total trip expenditures: a case study of instrumental variable (iv) regression analysis. *Tourism Economics*, 21(2):357–367(11).

Thrane, C. (2016a). Modelling tourists length of stay: A call for a back-to-basicapproach. *Tourism Economics*, 22(6):1352–1366.

Thrane, C. (2016b). Students' summer tourism: Determinants of length of stay (los). *Tourism Management*, 54:178–184.

Wang, E., Little, B. B., and DelHomme-Little, B. A. (2012). Factors contributing to tourists' length of stay in dalian northeastern chinaa survival model analysis. *Tourism Management Perspectives*, 4:67–72.