Unconditional Quantile Regression Analysis of UK Inbound Tourist Expenditures*

Abhijit Sharma†[1], Richard Woodward[2] and Stefano Grillini[3]

[1,3] *University of Bradford School of Management, Emm Lane, Bradford, BD9 4JL, United Kingdom.*

[2] *School of Strategy and Leadership, Coventry University, Coventry, CV1 5DL, United Kingdom.*

Ootober 2019

Abstract

Using International Passenger Survey (2017) data, this paper employs unconditional quantile regression (UQR) to analyse the determinants of tourist expenditure amongst inbound tourists to the United Kingdom. UQR allows us to estimate heterogeneous effects at any quantile of the distribution of the dependent variable. It overcomes the econometric limitations of ordinary least squares and quantile regression based estimates typically used to investigate tourism expenditures. However, our results reveal that the effects of our explanatory variables change across the distribution of tourist expenditure. This has important implications for those tasked with devising policies to enhance the UK's tourist flows and expenditures.

Keywords: tourist expenditures, unconditional quantile regressions, United Kingdom

JEL Classifications: C52, C83, Z30

*We would like to thank the editor of *Economics Letters*, Professor Badi Baltagi, and an anonymous referee for their helpful comments which have helped us to improve the paper. We are responsible for all remaining errors.

†Corresponding author. Email: A.Sharma12@bradford.ac.uk

1 Introduction

Boosting inbound tourist expenditures is an important government policy goal. There is a need for statistical analysis to accurately assess determinants of tourist expenditures. This is clearly illustrated for the UK where, in 2017, 39.2 million overseas visitors spent £24.5 billion (VisitBritain, 2017), making it the fifth largest recipient of international tourist revenues worldwide (UNWTO, 2018). Using data from the 2017 International Passenger Survey (IPS, ONS (2018)), this paper analyses the relationship between socioeconomic, demographic and satisfaction based drivers of tourism demand and expenditure amongst inbound tourists to the UK.

Existing studies (Thrane, 2014; Brida and Scuderi, 2013) establish the existence of statistically significant relationships between tourist expenditures and socio-economic, demographic and satisfaction variables. Here we use unconditional quantile regression (UQR) to provide more robust inference than OLS Firpo et al. (2009). QR provides estimates of effects along the conditional distribution of the dependent variable (i.e. its distribution when covariates are held constant). In other words, QR does not provide estimates of effects along the unconditional distribution of the dependent variable. Our results using UQR (see Hossain et al. (2018)) reveal how the effects of explanatory variables vary across the expenditure range with important implications for tourism promotion policies.

The IPS dataset we employ provides demographic details, information on expenditure and the length and purpose of stay. IPS (2017) also records visitor satisfaction, captured by the likelihood of recommending the UK as a tourist destination and how welcome visitors felt. IPS (2017) records stay as length of days. We augment IPS data by including further covariates from the gravity literature including visa requirements, population weighted distance, colonial relationships, and GDP per capita (Morley et al., 2014) thereby providing an appropriate model specification and addressing data limitations within previous studies (Thrane, 2014).

Table 1: Regression results - OLS, QR and UQR

	OLS	QR	UQR	QR	UQR	QR	UQR	QR	UQR	QR	UQR
Dep: Log expenditure		τ=10		τ=25		τ=50		τ=75		τ=90	
Stay (Log)	0.506***	0.405***	0.485***	0.482***	0.4274***	0.544***	0.497***	0.582***	0.537***	0.561***	0.557***
Male	0.112***	0.144***	0.165***	0.104***	0.069***	0.092***	0.103***	0.085***	0.139***	0.084***	0.087***
Age 0-24	-0.097**	-0.099	0.004	-0.056	-0.010	-0.030	-0.241***	-0.210***	-0.275***	-0.229***	-0.119**
Age 25-64	0.165***	0.284***	0.365***	0.189***	0.203***	0.114***	0.095***	0.001	-0.001	0.017	-0.001
Purpose: Holiday	-0.133***	0.108	0.725***	-0.042	0.278***	-0.089**	-0.139***	-0.231***	-0.408***	-0.341***	-0.706***
Purpose: Business	0.475***	0.481***	1.216***	0.613***	0.797***	0.559***	0.465***	0.409***	0.303***	0.262***	0.110*
Nationality: EU	-0.320***	-0.115	0.020	-0.175***	-0.078	-0.329***	-0.367***	-0.357***	-0.511***	-0.471***	-0.564***
Nationality: Europe non-EU	-0.179***	0.035	0.396***	-0.028	0.197***	-0.230***	-0.027	-0.239***	-0.453***	-0.339***	-0.925***
Visa required†	0.735***	0.766***	0.502***	0.733***	0.451***	0.710***	0.534***	0.643***	0.856***	0.852***	1.239***
Log GDP per capita∗	0.279***	0.406***	0.326***	0.301***	0.207***	0.256***	0.201***	0.206***	0.252***	0.194***	0.358***
Tourist activities	0.115***	0.193***	0.218***	0.141***	0.117***	0.093***	0.096***	0.076***	0.096***	0.063***	0.093***
Common language‡	-0.519***	-0.758***	-0.357***	-0.585***	-0.269***	-0.477***	-0.333***	-0.458***	-0.559***	-0.408***	-1.065***
Weighted distance‡	0.091***	0.170***	0.102**	0.102***	0.0978***	0.064***	0.099***	0.068***	0.103***	0.058***	0.051*
Former colony‡	0.243***	0.280*	-0.046	0.299***	0.017	0.239***	0.122**	0.303***	0.358***	0.250***	0.769***
Common currency‡	-0.709***	-1.810***	-1.381***	-0.671***	-0.519***	-0.568***	-0.432***	-0.375***	-0.336***	-0.296***	-0.434***
Felt welcome	0.042**	0.033	0.071	0.076***	0.0799***	0.0563***	0.0470***	0.035**	0.003	-0.017	0.000
Would recommend	0.069***	0.082*	0.126***	0.117***	0.0888***	0.079***	0.057***	0.045***	0.051***	0.013	0.044*
Constant	1.129***	-2.176***	-1.862**	-0.267	0.660**	1.443***	1.971***	2.892***	2.570***	4.160***	2.727***
Observations	19,534	19,534	19,534	19,534	19,534	19,534	19,534	19,534	19,534	19,534	19,534

OLS coefficients are estimated using robust standard errors. Each regression quantile is shown by value of τ at which R and UQR models are estimated. For UQR models, cluster robust standard errors are employed. Significance levels are: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. All data is sourced from the UK IPS, except ‡ from CEPII Gravity Data Set, ∗ from World Development Indicators 2018 and † from UK Foreign and Commonwealth Office and own calculations.

2 Methodology

We follow Koenker and Bassett Jr (1978) by generalising the optimisation procedure for a certain quantile of interest, τ, as follows:

$$\min_{\beta \in \mathbb{R}} \left[\sum_{t \in (t: y \geq x_t \beta)} \tau |y_t - x_t \beta_\tau| + \sum_{t \in (t: y < x_t \beta)} (1 - \tau) |y_t - x_t \beta_\tau| \right] \tag{1}$$

where $y_t - \beta x_t$ represents the residual from the regression of the vector of covariates, x_t, on the explanatory variable y_t. In addition to QR and OLS, we implement Firpo et al. (2009) type Unconditional Quantile Regressions (UQR) by computing a recentered influence function (RIF) which is constructed without reference to the covariates. The RIF is subsequently regressed on the explanatory variables:

$$RIF(y, \nu) = \nu(F) + IF(y, \nu) \tag{2}$$

where F is the cumulative distribution function of the response variable y and $\nu(F)$ represents the marginal effect on the parameter of the distribution F, when an observation is added or removed. In our paper, y corresponds to the logarithm of expenditure by tourists. The influence function (IF) can be defined as:

$$IF(y, \nu(F)) = \lim_{\Theta \to 0} \frac{[\nu((1-\Theta)F + \Theta\delta_y) - \nu(F)]}{\Theta} \tag{3}$$

These RIF values are subsequently used to perform a further regression on the covariates.

3 Empirical Results

We begin by using OLS to assess causal relationships. OLS based linear regression analysis mainly focuses on the mean, thereby neglecting the differing impact of particular variables across the expenditure range. In other words, the impact of variables such as visa and satisfaction varies depending on where the respondents are on the expenditure distribution, which OLS typically omits. In contrast, QR techniques do not provide estimates of effects along the unconditional distribution of the dependent variable. Further, in QR parameter distributions depend on the covariates chosen. As a consequence, the significance, size and direction of impact changes significantly as model specifications change. Using UQR addresses this limitation whilst retaining QR's methodological advantages. In particular, UQR allows us to reliably analyse the impact of changes in the distribution of explanatory variables on quantiles of the unconditional (or marginal) distribution of our response (or outcome) variable. Table 1, second column reports OLS estimates, while the remaining columns report QR and UQR coefficients for τ =10%, 25%, 75% and 90%. Our dependent variable is the logarithm of tourist expenditures. UQR estimates are obtained using bootstrapped standard errors from 1000 repetitions, which is well in excess of 200 repetitions used in Firpo et al. (2009). OLS regressions are significant for all our covariates but there are significant differences in estimates obtained using QR and UQR.

We exploit benefits arising from our methodological contribution for other covariates reported in Table 1. Notably, the UQR results reveal relationships that are underplayed or overlooked by OLS and QR. OLS indicates a significant coefficient of 0.279 for GDP per capita. QR shows a secular decline from τ=10 to 90 and UQR from τ =10 to 50. However, UQR rises sharply between τ =75 and 90 with a coefficient of 0.358 at τ =90. In other words, with higher GDP per capita levels we observe significant, higher tourist expenditures between τ=75 and 90, which is an effect discounted by OLS and belied by QR. Likewise, for visa requirements, OLS indicates a significant coefficient of 0.735, while QR estimates decline monotonically from τ=10 to 75, and then rise for τ=90. UQR estimates rise significantly at τ=75 reaching a highest absolute value of 1.239 at τ=90, which is significantly above the OLS value of 0.735, as well as QR and UQR values at the median (0.710 and 0.534, respectively). For instance, figure 1 shows that sections of the UQR curve (solid black line) at both ends of the distribution lie outside the 95% confidence intervals of the OLS coefficient. We clearly infer that the effect of visa requirement is is larger at the top end of the distribution, which was not evident by using QR. An important caveat to this conclusion is that selection effects could lead to some wealthy tourists, who are likely to spend more in relation to Visa requirement, appearing in our sample. Probing this aspect further, for example by making use of data on determinants of visa applications made by tourists (from a number of countries) is precluded by severe data limitations, not least resulting from confidentiality surrounding visa applications.

In short, UQR estimates report a much higher marginal effect of visa requirements on overall expenditures at the top end of the expenditure distribution, showing that OLS estimates significantly underreport effects. This may have paradoxical tourism policy implications. Although our results suggest that tourists visiting the UK who need a visa tend to spend more, the additional costs and frictions associated with visas may be harming expenditures by deterring potential inbound tourists. The UK may wish to consider easing visa requirements for countries with large numbers of wealthy outbound tourists.

The impact on expenditures of tourists from former colonies shows a coefficient of 0.243 using OLS, while QR shows significant effects ranging between 0.239 and 0.303 across the distribution. However, UQR indicates an insignificant effect for τ=10 and 25, and a small, significant effect at the median. There are much higher effects at τ=75 and 90, with a significant coefficient of 0.769 at τ=90, which is thrice as large as the OLS coefficient. These results suggest that to enhance tourist expenditures and revenues, the UK's policy makers should pay special attention to tourists from former colonies, particularly tourists with high disposable incomes. One unexpected finding is a negative coefficient for the common currency and common language dummy. By reducing transaction costs these two factors would ordinarily be expected to lead to a positive expenditure effect (see, e.g., Glick and Rose, 2002). Possibly this result arises because tourist expenditures, unlike other goods, are inseparable from the place where these services are consumed. Similarly, the widespread use of English as a second language may distort the impact of the language variable.

Visitor satisfaction is a key driver of tourist expenditures, not least because satisfied tourists are more likely to return and recommend that destination to others (Alegre and Garau, 2010). "Would recommend", is positive throughout and significant except for QR at τ=90. The OLS coefficient is 0.069, while QR rises until roughly the 30^{th} quantile and then declines monotonically. However, as Figure 2 shows, UQR estimates peak at τ=10 and decline until τ=60, after which they largely hold firm. What this shows is that UQR coefficients lie above the OLS coefficient (the middle horizontal line in Figure 2) until roughly τ=45 and lie below it from then on. There is clearly a greater impact on tourist expenditures of the likelihood of recommending the UK as a tourist destination towards the left of the distribution, as compared to the right. This suggests that tourists on lower incomes who view the UK more favourably and recommend visiting the UK are likely to have a greater impact on tourist expenditures, as compared to tourists on higher incomes, assuming tourist expenditures and tourist disposable incomes are positively related. The coefficients, despite being positive and significant, are smaller in magnitude towards the right of the distribution where we also observe larger disposable incomes and larger absolute expenditures. This implies that tourists on higher incomes and therefore undertaking higher expenditures on their visits are less likely to be influenced by the positive recommendation to visit the UK.

Our other satisfaction variable, "felt welcome", is significant and positive for the OLS model. For QR and UQR, welcome is insignificant at both extreme points, τ=10 and τ=90, while it is also insignificant for UQR at τ=75 (Figure 3). QR coefficients are positive and significant for τ=25, 50 and 75, while UQR is positive and significant for τ=25 and 50. In contrast to OLS, our quantile results indicate that the degree of welcome perceived is insignificant for tourist expenditures at both extreme tails of the distribution. UQR results indicate that higher and lower spenders (τ=10, 75 and 90) are unaffected by degree of welcome. This contributes to a positive impact on expenditures.

4 Conclusions

This paper assesses the impact of socioeconomic, demographic and satisfaction based variables on UK inbound tourist expenditures. We make a methodological contribution through our use of UQR and augment IPS data using important covariates from the gravity literature. This note highlights the value of UQR in addressing the inherent limitations within the OLS and QR methodologies used traditionally to analyse tourist expenditures. Specifically, the UQR methodology reveals the changing impact of independent variables across the expenditure distribution. UQR helps identify effects that are ignored or not evident using OLS and quantile regression analysis, with important tourism policy implications. For example, especially in the febrile macroeconomic and policy environment following the Brexit referendum, the UK should be cognisant of factors such as visa requirements which UQR reveals has an outsize impact on overall tourist expenditures.

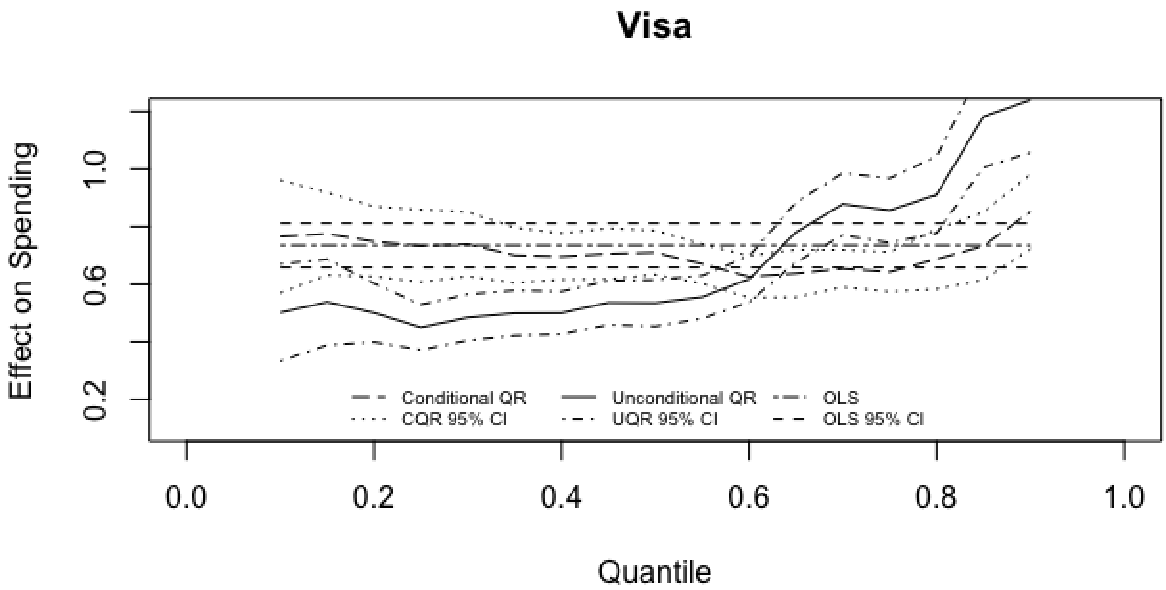

Figure 1: Effect of `visa requirement` on expenditures

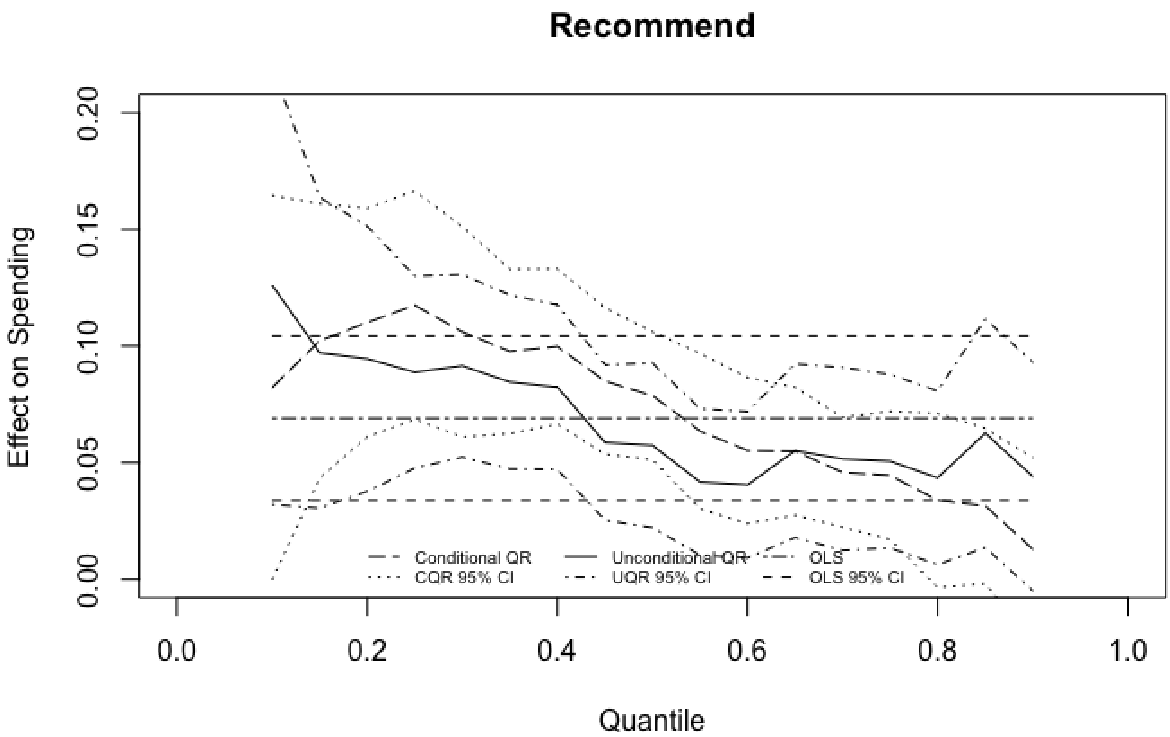

Figure 2: Effect of `recommend` on expenditures

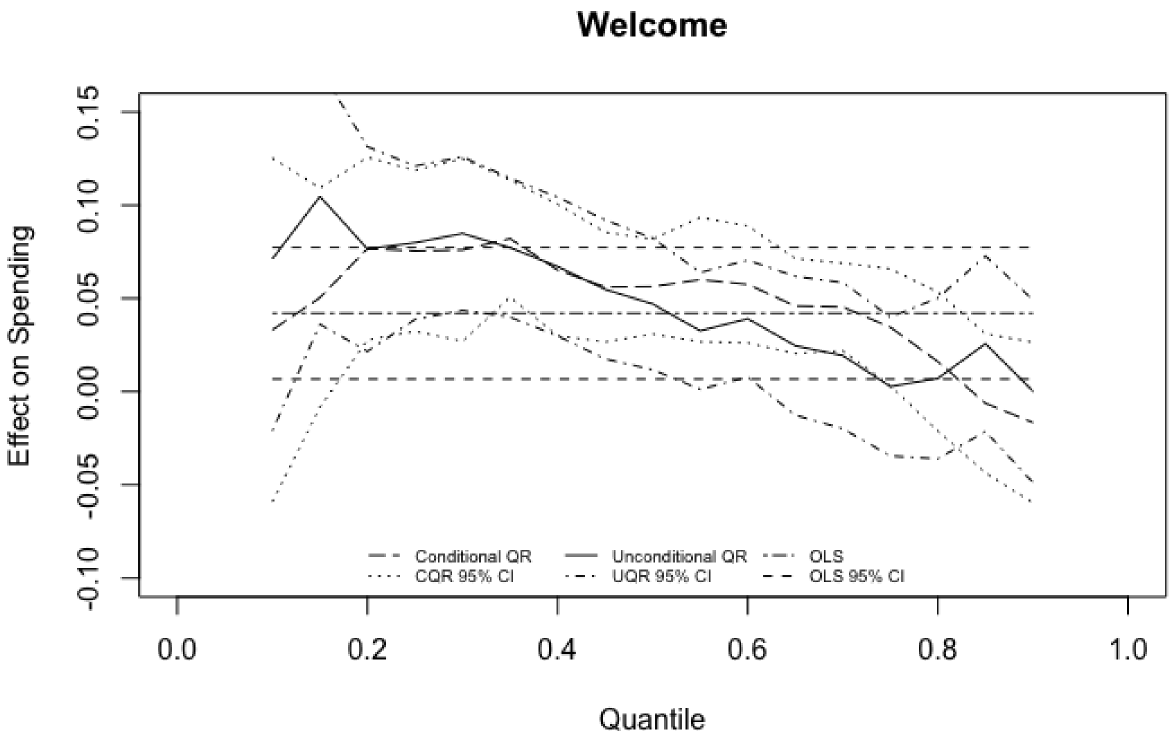

Figure 3: Effect of `welcome` on expenditures

References

Alegre, J. and Garau, J. (2010). Tourist satisfaction and dissatisfaction. Annals of Tourism Research, 37(1):52–73.

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

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

Glick, R. and Rose, A. K. (2002). Does a currency union affect trade? The time-series evidence. European Economic Review, 46(6):1125–1151.

Hossain, I., Saqib, N. U., and Haq, M. M. (2018). Scale heterogeneity in discrete choice experiment: An application of generalized mixed logit model in air travel choice . Economics Letters, 172:85–88.

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

Morley, C., Rosselló, J., and Santana-Gallego, M. (2014). Gravity models for tourism demand: theory and use. Annals of Tourism Research, 48:1–10.

ONS (2018). International Passenger Survey, 2017, Office for National Statistics [data collection]. 3rd edition, UK Data Service. SN: 8286.

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.

UNWTO (2018). Tourism Highlights 2018 Edition. `https://www.e-unwto.org/doi/pdf/10.18111/9789284419876`.

VisitBritain (2017). Inbound tourism trends by market. `https://www.visitbritain.org/inbound-tourism-trends`.