ESTIMATING PERCEPTIONS OF CONSUMERS USING AI: A CASE STUDY FOR GERMAN AND UK TRAVELERS

Onete Cristian Bogdan[1], Kejo Starosta[2], Budz Sonia[3] and Krutwig Michael[4]
[1) 2) 3) 4)] *The Bucharest University of Economic Studies*
Email: bogdan.onete@com.ase.ro; Email: k@ksfx.de;
Email: sonia.budz@yahoo.com; Email: michael@krutwig.com

Abstract
This empirical study analyzes the different perceptions of popular European tourist destinations in German and UK media reports. There are huge differences in the behavior of travelers from different countries, their preferred tourist destinations, accommodations, etc. Most of these differences can be explained by the differences in disposable income, exchange rates, gravity models that simply account for the travel distance, and some more. While all these metrics are reasonable explanatory variables, this study tries to elaborate whether there are further differences that originate from the general perception of a country.
To measure the differences in the perceptions, we analyzed the differences in the sentiments of the online media in Germany and UK toward potential travelling destinations and compared this information with tourist arrivals in these countries. The sentiments in the media were measured by artificial neural network software that analyzes the mood of the mass media.
The results revealed that there were indeed different perceptions in the German and UK online media for some European tourist destinations and these differences were reflected in tourist arrivals. These results suggested that there were either news on different topics or that the media in Germany and UK valued the same topics in different ways.
This information is useful for businesses in the tourism and hospitality sectors as well as for country promoters and news provider businesses to investigate the source and the impact of the information gap to optimize their operations.

Keywords
Tourism; online media; sentiment analysis; machine learning

JEL Classification
C45, C53, L83

Introduction

Besides several other differences among European countries, travelers from different countries prefer different travel destinations. Eurostat and the World Tourism Organization [WTO] track the behavior and the routes of travelers closely and provide timely data of tourism flows that show the differences between the countries. There are a large number of studies and models that try to explain the behavior of travelers in different countries (see e.g. Song and Li 2008, Dogru et al. 2017, Li et al. 2017, Dinis et al. 2017, and many more). Most of these models explain the differences by means of disposable incomes, exchange rates, gravity models that simply account for the travel distance, and some more factors (see e.g. Song & Li 2008). Further studies investigate the historical reasons for the differences in tourism flows such as immigration and refugee flows in the past (see e.g. Seetaram 2012, Moufakkir 2014).

While all these studies have reasonable explanations for the differences in tourism flows, the question still remains as to whether a fraction of these differences is solely attributed to the perceptions of travelers from different countries on travelling destinations. In a previous study, we showed that there are very strong correlations between the news sentiments in the German-speaking media and the tourist arrivals in popular tourist destinations in Europe (Starosta et al. 2018a). Moreover, we showed that these correlations are dependent on the arousal among the travelers created by the news reports (Starosta et al. 2018b). This study also suggests that it is possible to classify which types of news reports lead to stronger or weaker arousal levels. These findings directly lead to the question of whether news reporting in different countries of origin differs and there are also differences in the sentiments or if all further differences in the behavior of tourists can be attributed to the classical models that neglect or overlook the perceptions of travelers.

A comparison between Germany and UK seems reasonable as both have a similar median disposable income (Germany $25,140, UK $21,576)* and the role of the gravity is less important as their distance is not too far away (Avg. distance ~800km). Further, both are western European countries that are not connected to the Mediterranean Sea and they have some common tourist destinations.

To cost-effectively measure the perceptions of German and UK travelers to European tourist destinations, we measure the sentiments toward these countries in the German and English media. While this not a direct measure—such as a survey—it is probably one of the best proxies that could be used to analyze the sentiment of a country toward another country. Also, we suppose that the trends of specific perceptions of a country toward another country are mutually reinforcing between the news reports in the media and the general perceptions of the citizens.

We analyze the sentiments in the German and UK media with artificial neural network software that is able to classify the news reports of very comprehensive news streams in both countries. The software rates every news item, either positive or negative, that refers to a potential tourist destination. This information can be used to generate time series that reflect the sentiments in the German and UK media over time. Each pair of origin and destination (e.g. origin Germany and destination Italy) gets two time series: one with the number of positive news and one with the number of negative news. We compare the correlations in these time series between Germany and UK to see the differences in perceptions and also between their corresponding tourist arrivals to see if the different perceptions come along with different behaviors in arrivals and different preferences for travel destinations.

Theoretical Foundations

Starosta et al. (2018a) analyzed the impact of online media on tourist arrivals in Europe. While this study was able to show strong correlations between the sentiment in the German-speaking media and tourist arrivals in Europe, further research showed that this relation might be different for other countries and language areas. We expect that, besides many other factors, perceptions play an important role. Many studies suggest that the reporting by the mass media drives the perceptions, especially those pertaining to international relations. Wanta et al. (2004), McComb and Reynolds (2002), and others show that the agenda is set by the mass media. Furthermore, Wanta et al. (2004) argue that the more negative coverage

* Source OECD http://stats.oecd.org/Index.aspx?DataSetCode=IDD

a nation receives, the more likely the respondents think negatively about the nation. Brewer et al. (2003) find similar evidence that a "news story that presents a frame linking an issue to a foreign nation in a way that suggests a particular evaluative implication may shape how audience members judge that nation." Zeitzoff (2016) analyzes how social media can influence armed conflicts. Apart from the studies on media influence on international relations, a tremendous number of different studies focus on several other topics of media influence.
Besides Starosta et al. (2018a) and the studies that analyze media influence on different groups, tourism research provides the foundations and the explanatory variables that do not account for the perceptions of travelers. Song and Li (2008) provide a comprehensive overview of tourism demand forecasting between 2000 and 2008, which summarizes most of the classical indicators based on econometric and time series models. A recent study by Dogru et al. (2017) enhances the classical tourism demand models. Some newer studies in the area of tourism research also apply methods that engage big data from the web, such as search engine query volumes and sentiments from social media posts. Li et al. (2017) propose a method to forecast tourism demand via a composite search index built with the help of search engine query volumes. In a similar study, Dinis et al. (2017) analyze the correlations between the search query volumes of specific query terms for different regions in Portugal and the overnight stays spent in these areas. Mathieson and Wall (1982) identify five steps in the decision-making process of tourists—identification of needs, information search, alternatives evaluation, choice and purchase, and post-purchase evaluation. We expect the active search process—as measured by the search engine query volumes—to start at a later step than the opinion-forming for a potential destination; thus, we expect the approach of sentiment analysis in the mass media to have a longer lead to the actual arrivals than the query volumes. Although there are a large number of sentiment analysis studies in the area of tourism research, they mostly focus on microeconomic data and the analysis of travel reviews. Ye et al. (2009) analyze the sentiments about travel destinations in online reviews with the help of different machine-learning approaches and evaluate their qualities. Gonzales-Rodriguez et al. (2016) analyze the post-visit and pre-visit tourist destination images through electronic word-of-mouth sentiment analysis and the perceived helpfulness of reviews on the Ciao platform. Ren and Hong (2017) present a very interesting analysis on microeconomic data. They not only analyze the sentiments in online reviews but also attempt to determine the sentiment toward a specific topic within the review. Chaabani et al. (2017) use sentiment analysis to track touristic reviews in social media; this is one of the very few analyses to also focus on the macroeconomic structures of tourism. They analyze the tourists' viewpoints of Tunisia as a travel destination after Arab Spring. Yuan et al. (2016) uses the dataset "The Global Data on Events, Location, and Tone" to model the image of China in the mass media and how its relations have evolved with time to model tourism demand for the country. Apart from the demand modeling approaches, there are other studies that analyze the impact of shocks and instability on tourism destinations. Sönmez and Graefe (1998) give a comprehensive overview of the studies on tourism, terrorism, and political instability. In addition to the studies focusing on the tourism industry, Kotler and Gertner (2002) treat a country as a brand or product and theoretically analyze how images affect attitudes toward a country's products and services and its ability to attract investments, businesses, and tourists.

The domain of computer science provides the foundations for sentiment analysis with the help of machine-learning algorithms. Computational sentiment analysis is a research area that has been growing rapidly since the 2000s; the first comprehensive studies were conducted by Pang and Lee (2008). Sentiment analysis can be done with any machine-learning algorithm, such as Naïve Bayes, support vector machines, logistic regression,

conditional random fields, artificial neural networks, etc. Moreover, there are less generic approaches that try to detect the sentiment with the help of wordlists, n-gram lists, and similar methods. The latest comprehensive study on sentiment analysis with the help of machine learning is by Liu (Liu). While there is a huge volume of research in the field of sentiment analysis and machine learning, this study focuses on production-ready and applicable algorithms. Several studies show that even simple algorithms usually perform well enough for this kind of study, with more than 85% of correct sentiment predictions (Domingos & Pazzani 1997, Endres 2003, Potts 2011). In the domain of applied computer science, there have been various efforts to analyze the impact of sentiments on stock and financial markets. Zhang et al. (2010) report correlations between the emotional words Hope, Happy, Fear, Worry, and others and the indices of Dow Jones Industrial, NASDAQ, S&P500, and VIX. When a lot of hope or fear is expressed on Twitter, the indices tend to fall the next day. Bollen et al. (2011) optimize this approach and attempt to forecast the Dow Jones Industrial index with the help of Twitter messages (Tweets). Zhang and Skiena (2010) successfully forecast single stocks with the help of sentiment analysis. Feldman et al. (2011) analyze economic news with an algorithm that provides a deeper textual understanding. Si et al. (2013) implement an approach that tries to analyze the sentiment on specific aspects and topics, similar to Ren and Hong (2017).
Finally, some studies from the fields of economics and finance use mass media sentiment analysis to forecast or nowcast macroeconomic data. While these studies are from different domains, their methodology is more similar to this study than most of the studies in the domain of tourism. Daas and Puts (2014) analyze the relation between Dutch messages in different social media channels and the Dutch consumer confidence index. Förschler and Alfano (2017) show correlations between financial news, the leading German indices, and the number of incoming orders. Shapiro et al. (2017) from the Federal Reserve Bank of San Francisco focus on forecasting business cycle indicators and show that their proprietary news sentiment index is strongly correlated with some of them. This study is based on the same concepts as those in the studies by Daas and Puts (2014), Förschler and Alfano (2017), and Shapiro et al. (2017), which we have implemented in Starosta (2018a).

Research Question

This study tries to clarify the question as to whether the differences between Germany and UK pertaining to favorite tourist destinations and the differences in arrivals are solely based on classical and historical assumptions, such as distance to destination, disposable income, past immigration flows, and explanatory variables that are found in the current literature (e.g. as summarized by Song & Li 2008) or if there are more differences in perceptions between the travelers from Germany and UK that contribute to the differences in tourist flows.
The null hypothesis is that there are no further contributions that are based on travelers' perceptions and that there should be no big differences in perceptions between the travelers from different countries.

H0: There are no further differences between the perceptions of travelers from Germany and UK, and tourism flows can be solely explained by the existing explanatory variables.

This hypothesis is reasonable because the media in both countries in general state that it is objective, that the most important news are delivered in a timely manner, and that no news are deliberately kept secret. This should lead to very similar news reporting in Germany and UK, and so the perceptions in these countries should not differ too much.

Methodology

The study is divided into several steps of data acquisition and data analysis. The details of the methodology are explained in Starosta et al. 2018a. In contrast to the Starosta et al. 2018a study, this study focusses on the differences in the perceptions between Germany and UK rather than on the correlations between the perceptions and the tourist arrivals.

Results

Table 3 shows the results of the tourist destinations of German and UK travelers. The first column displays the destination and the second column displays the correlation between the sentiment indices of Germany and UK. While most of the perceptions of the countries correlate very strongly—which is the expected behavior if the news reporting is neutral and objective—the null hypothesis could be rejected for Austria, Turkey, Croatia, and Egypt (Note that the null hypothesis is rejected when the correlation is weak).

Table 3: Correlation of Sentiments in German and English Online Media

Destination	$Adj\ R^2$
Spain	0.629
France	0.803
Italy	0.820
USA	0.990
Austria	**0.348**
Turkey	**-0.011**
India	0.943
Croatia	**0.378**
Egypt	**0.497**
Greece	0.990
Portugal	0.975
Italy	0.806

Discussion

The results show that there are indeed different perceptions in Germany and UK for some countries and also that the media reports are different for the same events.
The biggest differences appear in the countries that also enjoy different popularities among travelers. Austria and Turkey, two very popular tourist destinations for German travelers that ranking 3rd and 4th, also belong to the countries for which the null hypothesis could be rejected. These countries do not appear among the top destinations for UK travelers at all. Austria is a country with very few British travelers, while Turkey is also an important tourist destination for British travelers, recording roughly one-third fewer British travelers than those Italy has.[†] Besides these very important countries for German travelers, the null hypotheses could be rejected for Croatia and Egypt as well.

In the case of Austria, the sentiment in the UK media was more or less neutral between 2009 and 2015, and then the sentiment took a negative turn in the aftermath of the European migrant crisis that started in 2015. The European migrant crisis affected Austria very badly because of its location on the Balkan route. The sentiment in the UK news persistently deteriorated until the end of the observation period. The sentiment in the German media was

[†] Source: EUROSTAT
http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tour_dem_tnw&lang=en

not at all influenced by the events that happened and the decisions that were made during the European migrant crisis in Austria. The sentiment in the German media toward Austria kept increasing positively. It seems that the different opinions on the refugee policies divided the perceptions in the media and thus, as shown by Seetaram, Moufakkir, and others, also the perceptions of potential travelers. As the sentiment in the UK media changed in a profound way and was very negative until the end of the observation period, there might also be further topics—e.g. those related to Brexit—which contributed to further differences in perceptions.

In comparison, the case of Turkey is different. While the news sentiment in the German media almost fitted perfectly with tourist arrivals ($Adj\ R^2$ 0.899) [‡], the sentiment in the UK media was negative during the whole observation period. Turkey's autocratic shift, which happened before and after the coup attempt in July 2016, had a significant impact on news reporting in Germany with heavily deteriorating sentiment in German media. In contrast, the coup attempt and the autocratic shift in Turkey did not influence the sentiment in the UK media at all, since the reporting on Turkey was very negative anyway. One reason for these results might be the large number of citizens with Turkish origin in Germany, accounting for around 4% of the German population. In 2015, around 2,851,000 citizens with Turkish origin lived in Germany.[§] Therefore, we expect that the results for Germany are hardly biased by its citizens with Turkish origin. Still, it is interesting to see the differences in perceptions. In a world with unbiased and objective news, the perception of a country should not depend on the origins of its citizens.

The case of Croatia is very similar to that of Turkey. The perception of the German media was positive for almost the entire observation period. Only for a short span in 2015, when Croatia had closed its borders to Serbia and a large number of refugees crossed into Croatia because of the European migrant crisis, the sentiment in the German media worsened. In contrast, the sentiment in the UK media kept on worsening all the time.

Finally yet importantly, in the case of Egypt, the media in Germany reacted on every event that happened during the Arab Spring: the overthrow of President Mubarak in February 2011, the overthrow of President Mursi in July 2013, and the bombing of a flight in October 2015. In contrast, the sentiment in the UK media kept on worsening all the time and the decline in sentiment did not speed up with special events. Interestingly, the wobbling sentiment based on events is not unique to the German media in general. Rather, this is a fact pertaining to this specific case. The perceptions regarding other countries also show wobbling media (and sentiment) reactions for every event in UK news reporting.

While there is a broad consensus in the perceptions for most countries, the perceptions for specific countries can differ greatly. The root cause for these different perceptions cannot be elaborated in the course of this study. Especially the case of Austria suggests that one reason for the different perceptions might be different opinions on the policies implemented by a country—e.g. the refugee policy in Austria.

Conclusions

This study showed how the perceptions in Germany and UK differed for two of their most important tourist destinations and there are more differences for other destinations as well. Like the differences in perceptions, the consensus for other destinations is equally interesting. For instance the media perceptions about the US or India are the same in both German and UK media.

[‡] Source: Starosta et al. 2018a

[§] Source: Migration Report of the BAMF: https://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/Migrationsberichte/migrationsbericht-2015.pdf?__blob=publicationFile

We could show that our perceptions, which are driven by media reports and by the agenda that the media sets for us, have a big impact on many aspects for tourism businesses and other sectors.

References/Bibliography

Bollen, J., Mao, H., Zeng, X., 2011. *Twitter Mood Predicts the Stock Market*, in Journal of Computational Science, Vol. 2, No. 1, March 2011, pp. 1–8, https://doi.org/10.1016/j.jocs.2010.12.007

Brewer, P.R., Graf, J., Willnat, L., 2003. *Priming or Framing: Media Influence on Attitudes toward Foreign Countries*, https://doi.org/10.1177/0016549203065006005

Chaabani, Y., Toujani, R., Akaichi, J., 2017. *Sentiment Analysis Method for Tracking Touristic Reviews in Social Media Network*, In: De Pietro G., Gallo L., Howlett R., Jain L. (eds) Intelligent Interactive Multimedia Systems and Services 2017. KES-IIMSS 2017. Smart Innovation, Systems and Technologies, Vol. 76. Springer, Cham, https://doi.org/10.1007/978-3-319-59480-4_30

Daas, P., Puts, M., 2014. *Social Media Sentiment and Consumer Confidence*, in ECB Statistics Paper 5, 2014

Dinis, M. G. F., Costa, C. M. M., Pacheco, Osvaldo, M. R., 2017. *Similarities and Correlation between Resident Tourist Overnights and Google Trends Information in Portugal and Its Tourism Regions*, in Tourism & Management Studies, Vol. 13, No. 3 (2017), http://www.tmstudies.net/index.php/ectms/article/view/940 (2017-10-30 20:44)

Dogru, T., Sirakaya-Turk, E., Crouch, G. I., 2017. *Remodeling International Tourism Demand: Old Theory and New Evidence*, in Tourism Management, Vol. 60, June 2017, pp. 47–55, https://doi.org/10.1016/j.tourman.2016.11.010

Domingos, P., Pazzani, M. J., 1997, *On the Optimality of the Simple Bayesian Classifier under Zero-one Loss*, http://www.ics.uci.edu/~pazzani/Publications/mlj97-pedro.pdf (12.10.2015 22:43)

Endres, J., 2003. *Spam oder nicht Spam? E-Mail sortieren mit Bayes Filtern*, c't 17/2003, p. 150

Feldman, R., Rosenfeld, B., Bar-Haim, R., Fresko, M., 2011. *The Stock Sonar—Sentiment Analysis of Stocks Based on a Hybrid Approach*, in Proceedings of the Twenty-Third Innovative Applications of Artificial Intelligence Conference, AAAI Press, Menlo Park 2011, pp. 1642–1647

Förschler, F., Alfano, S., 2017. *Reading between the Lines: The Effect of Language Sentiment on Economic Indicators*, in: Enterprise Applications, Markets and Services in the Finance Industry, Springer Press, Cham 2017, pp. 89–104

González-Rodríguez, M. R., Martinez-Torres, R., Toral, S., 2016. *Post-visit and Pre-visit Tourist Destination Image through eWOM Sentiment Analysis and Perceived Helpfulness*, in International Journal of Contemporary Hospitality Management, Vol. 28, No. 11, pp. 2609–2627, https://doi.org/10.1108/IJCHM-02-2015-0057

Kotler, P., Gertner, D., 2002. *Country as Brand, Product, and Beyond: A Place Marketing and Brand Management Perspective*, in Journal of Brand Management (2002), Vol. 9, pp. 249–261, https://doi.org/10.1057/palgrave.bm.2540076

Li, X., Pan, B., Law, R., Huang, X., 2017. *Forecasting Tourism Demand with Composite Search Index*, in Tourism Management, Vol. 59, April 2017, pp. 57–66, https://doi.org/10.1016/j.tourman.2016.07.005

Liu, B., 2015. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, Cambridge University Press, 1st Edition, Cambridge 2015, 978-1107017894

Mathieson, A., Wall, G., 1982. *Tourism: Economic, Physical, and Social Impacts*

Moufakkir, O., 2014, *What's Immigration Got to Do with It? Immigrant Animosity and Its Effects on Tourism*, https://doi.org/10.1016/j.annals.2014.08.008

McCombs, M., Reynolds, A., 2002. *News Influence on our Pictures of the World*. In Media Effects: Advances in Theory and Research, 2nd Edition, 2002, ISBN 0-8058-3864-3, pp. 1–19

Zeitzoff, T., 2016. *Does Social Media Influence Conflict? Evidence from the 2012 Gaza Conflict*, https://doi.org/10.1177/0022002716650925

Pang, B., Lee, L., 2008. *Opinion Mining and Sentiment Analysis*, in Foundations and Trends in Information Retrieval, Vol. 2, No. 1–2, pp. 1–135, http://dx.doi.org/10.1561/1500000011

Potts, C., 2011. *Sentiment Symposium Tutorial: Classifiers*, http://sentiment.christopherpotts.net/classifiers.html (23.02.2015, 14:07)

Ren, G., Hong, T., 2017. *Investigating Online Destination Images Using a Topic-Based Sentiment Analysis Approach*, in Sustainability 2017, http://dx.doi.org/10.3390/su9101765

Seetaram, N. 2012. *Immigration and International Inbound Tourism: Empirical Evidence from Australia*, https://doi.org/10.1016/j.tourman.2012.02.010

Shapiro, A., Sudhof, M., Wilson, D. J., 2017. *Measuring News Sentiment*, Federal Reserve Bank of San Francisco Working Paper 2017-01, http://www.frbsf.org/economic-research/publications/working-papers/wp2017-01.pdf (16.02.2017 17:51)

Si, J., Mukherjee, A., Liu, B.; Li, Q.; Li, H.; Deng, X., 2013. *Exploiting Topic based Twitter Sentiment for Stock Prediction*, in Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, http://www.aclweb.org/anthology/P13-2005

Sönmez, S.F., Graefe, A.R., 1998. *Influence of Terrorism Risk on Foreign Tourism Decisions*, in Annals of Tourism Research Vol. 25, No. 1, January 1998, pp. 112–144, https://doi.org/10.1016/S0160-7383(97)00072-8

Starosta, K., Budz, S. and Krutwig, M., 2018a. *The Impact of Online Mass Media on Tourist Arrivals in Europe*, Submitted

Starosta, K., Budz, S. and Krutwig, M., 2018b. *Artificial Neural-Network-Based Emotion Classification in the Online Media for Tourism Businesses*, in Proceedings of the 5th European Conference on Social Media, Limerick 2018

Song, H., Li, G, 2008. *Tourism Demand Modelling and Forecasting—A Review of Recent Research*, in Tourism Management Vol. 29, No. 2, April 2008, pp. 203–220, https://doi.org/10.1016/j.tourman.2007.07.016

Wanta, W., Golan, G., Lee, C., 2004. *Agenda Setting and International News: Media Influence on Public Perceptions of Foreign Nations*, https://doi.org/10.1177/107769900408100209

Ye, Q., Zhang, Z., Law, R., 2009. *Sentiment Classification of Online Reviews to Travel Destinations by Supervised Machine-Learning Approaches*, in Expert Systems with Applications, Vol. 36, No. 3, Part 2, April 2009, pp. 6527–6535, https://doi.org/10.1016/j.eswa.2008.07.035

Yuan, Y., Liu, Y., Wei, G., 2017. *Exploring Inter-country Connection in Mass Media: A Case Study of China,* in Computer, Environment and Urban Systems, Vol. 62, March 2017, pp. 86–96, https://doi.org/10.1016/j.compenvurbsys.2016.10.012

Zhang, X., Fuehres, H., Gloor, P. A., 2010. *Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear"*, in Procedia—Social and Behavioral Sciences Vol. 26, 2011, pp. 55–62, https://doi.org/10.1016/j.sbspro.2011.10.562

Zhang, W., Skiena, S., 2010. *Trading Strategies to Exploit Blog and News Sentiment*, in Fourth International AAAI Conference on Weblogs and Social Media, http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1529/1904 (17.10.2015 22:01)