Predicting international tourist arrivals in Greece with a novel sector-specific business leading indicator

Anastasiou, Dimitris and Drakos, Konstantinos and Kapopoulos, Panayotis

Athens University of Economics and Business, Alpha Bank, University of Piraeus

July 2022

Predicting international tourist arrivals in Greece with a novel sector-specific business leading indicator

Dimitris Anastasiou[†,1] (anastasioud@aueb.gr)
Konstantinos Drakos[†] (kdrakos@aueb.gr)
Panayotis Kapopoulos[1,2] (panos.kapopoulos@unipi.gr)

[†]Department of Accounting and Finance, Athens University of Economics and Business, Greece
[1]Economic Research Division, Alpha Bank, Greece
[2]University of Piraeus, Greece

Abstract

We introduce a novel tourism-specific business expectations sentiment index and explore whether it can operate as a leading indicator for international tourist arrivals in Greece. Using monthly data spanning 2002-2021 and employing a VAR model, we document that this newly introduced tourism-specific business expectations serves as a leading indicator, whose higher levels foreshadow increased demand for international travel. We also find that its inclusion in a tourism-oriented model increases forecasting accuracy, which can be utilized by travel agent businesses, local government officials and policymakers in their efforts to predict tourist arrivals in Greece.

Keywords**:** Leading Indicator; Expectations; Tourist Arrivals; VAR; Greece

JEL classification: C22, C53, Z30

Disclaimer: The views and opinions expressed in this paper are those of the authors and do not reflect their respective institutions.

Declaration of Competing Interest: No conflict of interest exists in the submission of this manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are not publicly available due to ethical reasons. However, all the data can be found from open sources, as described in the main manuscript. The analysis was conducted using the EViews 11 Software.

1. Introduction

A growing body of literature has focused on either the macroeconomic determinants of tourism demand or climate-confidence indicators to forecast tourist arrivals and receipts in several countries (see among others, Icoz *et al.*, 1998; Smeral and Weber, 2000; Kulendran and Witt, 2003; Song *et al.,* 2003; Papatheodorou and Song, 2005; Kim and Moosa, 2005; Athanasopoulos and Hyndman, 2008; Saayman and Saayman, 2008; Song *et al.,* 2010; Kim *et al.,* 2012; Tavares and Leitao, 2017; Santamaria and Filis, 2019).

Although prior literature suggests possible leading indicators for the tourism industry (Kulendran and Witt, 2003; Crotts *et al.,* 1993; Turner and Witt 2001; Guizzardi and Stacchini 2015; Gholipour and Tajaddini 2018; Gholipour and Foroughi, 2020; Yost *et al.,* 2020), none of them proposed a tourism-specific sentiment pertinent to the supply side of the market[1]. Instead, they examined generalized consumer and/or business confidence indicators as potential candidates serving as a leading indicator. In some more detail, Guizzardi and Stacchini (2015) were the first who examined whether supply-side soft information is effective in real time forecasting of hotel[2] arrivals. However, their analysis is limited and only related to hoteliers, not capturing the whole tourism-related business (including tour operators, travel agencies, etc). In addition, they examine its forecasting power only on tourist arrivals in the hotels of the province of Rimini in Italy. On the contrary, our tourism-specific business confidence is examined on the international tourist arrivals in Greece, including hotels, camping grounds, recreational vehicle parks, trailer parks for holidays, and other short-stay accommodation.

The purpose of this study is thus to examine whether the novel composite expectational leading indicator pertinent to the tourism industry (i.e., hotelliers, tour operators, travel agencies,

[1] Chen *et al.* (2021) were the first who constructed a resident-specific sentiment describing local residents' overall perceptions of and emotional dispositions toward a dominant tourist market.
[2] Choi (2003) developed an economic indicator system for the US hotel industry to project the industry's growth and turning points, while Lim *et al.* (2009) examined a variety of time series models to forecast both hotel and motel guest nights in New Zealand.

etc.) which we propose, can act as a leading indicator for tourism demand. The latter is proxied by the number of international tourist arrivals in Greece. Using monthly data over the period of 2002-2021, the results from the estimated Vector Auto Regressive (VAR) model show that higher levels of the above-mentioned tourism-specific leading indicator (composite business confidence) in Greece increase demand for international travels.

Apart from investigating the impact of expectations on tourist arrivals, this paper also provides a methodological contribution by utilizing impulse response function within a VAR framework. This involves simulating impulse response functions (IRFS) from the shock of the leading indicator to provide information on the size of the reaction and the duration of the effects on future tourist arrivals. Confidence bands are computed using 200 Monte Carlo Simulation to determine the statistical reliability of the response. Unlike previous studies examining confidence/leading indicators on tourism, we do not only use a leading indicator to explain the demand function of tourist arrivals. Instead, we generate shocks from the leading indicator through the impulse response function by utilizing the VAR and investigate the persistence of these shocks on international tourist arrivals.

Furthermore, we conduct a forecasting exercise and we find that the leading indicator also reduces forecast errors when it is incorporated in a tourist arrivals model. Accurate tourist arrival forecasts are important for policymakers because they may be used to make policy decisions, aimed at improving economic development, wellbeing and employment, especially in tourism destination countries like Greece (Song and Witt, 2006; Gounopoulos *et al.,* 2012). In addition, accurate forecasts are also important at industrial level (e.g. hotels, tour operators, airlines, etc.), as they allow firms to produce more accurate budgets (Hassani *et al.,* 2017).

Our study makes a significant contribution by introducing a new explanatory variable (expectational leading indicator) to international tourism demand modeling. In particular, our study contributes to the literature in a threefold manner. First, we construct for the first time a tourism-specific leading indicator based on business expectations for the whole tourism sector and investigate whether it can serve as a measure of tourism company managers' optimism or pessimism toward their near-future business performance. Second, while the relationship between consumer confidence and tourism has been established in previous studies (e.g., Crotts *et al.,* 1993; Turner and Witt 2001; Gholipour and Tajaddini 2018), the link between business-wise confidence indicators and international tourist arrivals has received relatively little research attention in the tourism literature (Guizzardi and Stacchini, 2015). Finally, we add on the growing research of how non-fundamental variables, such as sentiment, expectations and/or business confidence, affect general aspects of the economic environment (see among others, Kulendran and Witt, 2003; Baker and Wurgler, 2006; Guizzardi and Stacchini 2015; Alaei *et al.,* 2019; Fu *et al.,* 2019; Anastasiou and Katsafados, 2020; Hao *et al.,* 2020; Anastasiou and Drakos, 2021; Anastasiou *et al.,* 2021; Letdin *et al.,* 2021; Anastasiou *et al.,* 2022a; Anastasiou *et al.,* 2022b).

The rest of the paper unfolds as follows. Section 2 outlines the reasons why Greece consists of an ideal laboratory. Section 3 describes the data and the econometric approach we employed, while section 4 presents the empirical findings. Finally, Section 5 concludes.

2. Stylized Facts: Why does the Greek case matter?

Tourism is one of Greece's most important industries, having enormous multiplier effects on the country's economic activities, laying the path for long-term development. With the improvement of hotel amenities and the entry of chain hotels into the domestic market, the Greek hotel sector has continued to expand for over a decade, becoming Greece, one of the strongest and fastest-growing accommodation areas. The upgrade of infrastructure in the context of the Olympic Games in the 2000s, as well as the enhancement of services in other areas (food and beverage services), helped Greece fulfill rising foreign demand in the 2010s. Travel receipts in

Greece increased at a pace higher than the worldwide tourism increase mainly due to the significant improvement in the unit labour cost and price competitiveness associated with internal devaluation policies pursued in the context of the rescue programs, well known as Memorandum of Understandings (MoUs), as well as the geopolitical uncertainty of some competitors, especially after the “Arab Spring” episodes (Adamopoulou *et al.,* 2022). So, despite the prolonged recessionary shock after the global economic crisis in Greece, the sector’s performance remained strong since it was mostly driven by external demand.

In the pre-pandemic period, travel receipts were an important part of exports of services “covered” somehow, on average, from 2015 to 2019, the trade deficit by 76%. In other words, tourism activity was the main source of “financing” the deficit in trade balance shaping a tourism–led growth pattern[3]. So, in many ways, tourism served as a life craft for the Greek economy to get through the storm of the 2010’s economic crisis. However, the high dependence of the Greek economy on tourism makes it vulnerable to external shocks, such as, later, the pandemic crisis where as we have seen, the countries with the comparatively higher tourism contribution to the Gross Value Added suffered the largest losses in terms of GDP in during pandemic (Adamopoulou *et al.,* 2022).

In addition, the tourism industry in Greece is primarily reliant on international visitors. The percentage of nights spent by foreign tourists in tourist lodging businesses has increased from 69% to 84% in 2019. Furthermore, leisure was the primary motive for visitors visiting Greece, accounting for 94% of total revenues on average between 2010 and 2019. In terms of employment, Greece’s lodging and food services industries employed about one out of every ten people in 2019, the largest employment proportion in the EU-27. It is also worth mentioning that since 2008, this proportion has increased by 2.8% (7% of total employment).

[3] For an empirical verification of the existence of such a pattern, see Lolos *et al*, (2021).

Finally, in terms of the country's international performance, according to the World Economic Forum's Travel & Tourism Competitiveness Index, Greece was placed 25th out of 140 nations in 2019, up seven places from 2013. Furthermore, Greece's foreign visitor arrivals market share increased to 2.1 percent in 2019 from 1.6 percent in 2010. Spain, Portugal, and Croatia, the country's main European competitors in the Mediterranean, grew their shares in 2019 compared to 2010, while France, Turkey, and Italy dropped theirs.

All the above demonstrate the importance of the tourism sector in Greece, and hence, finding a proper leading indicator may further improve country's long-term growth dynamics. Overall, the tourism sector is one of the most important areas for Greece's economic development, accounting for a significant portion of its GDP and employment numbers. Given the significance of tourism and its rapid rise in recent years, having a leading indicator of foreign travel demand would be critical for tourist authorities and operators, macroeconomic policymakers, and airline executives.

3. Data, Variables and Methodology

This section describes the data, the variables and the econometric methodology considered in our study.

3.1 Data and Variables

In this study, we use monthly data from EUROSTAT[4] for international tourist arrivals (TOURIST) in Greece from 2002 to 2021. As international tourist arrivals, EUROSTAT defines the inbound tourists from foreign countries who visit a country (Greece in our case) and live in hotels, camping grounds, recreational vehicle parks, trailer parks for holidays, and other short-stay accommodation. The choice of this time span was primarily based on the data availability.

[4] Data on tourist arrivals is also available from the site of Bank of Greece. However, because they are only offered on a quarterly basis, they are not chosen for the purposes of our analysis.

In addition, during this period, as we have already described in Section 2, many extreme, either positive or negative, events occurred and significantly impacted international tourist arrivals in Greece.

An important feature of our study is that it includes the ongoing COVID-19 pandemic. International tourist arrivals in Greece declined by 76.5% in 2020 compared to 2019, whereas, travel receipts were reduced by 76.2%. It should be noted, however, that 2019 was a record year for the Greek tourism, as 31,4 million people visited the country (+4.1%, on an annual basis), corresponding to Euro 18,2 million receipts from tourism, (+18%, on an annual basis). In addition, travel receipts accounted for almost 1/4 of total exports in 2019, whereas this share dropped to only 8.2% in 2020. Thus, we are able to investigate the performance of the tourism-specific leading indicator not only during “normal times”, but also during extreme events.

In order to construct an expectational leading indicator reflecting the tourism sector business sentiment, we obtain data for tourism-specific business expectations from the European Commission’s (EC) harmonized survey program, managed by the Directorate-General for Economic and Financial Affairs (DG ECFIN). The survey data generated within the Joint Harmonised EU Programme of Business and Consumer Surveys framework are particularly useful for monitoring economic developments. The monthly services survey provides information about managers’ assessment of their future business situation. In particular, we take advantage of the services sub-sectors business expectations, and we consider the following two questions from the answers of which the EC then constructs two distinct indicators showing expectations for the tourism industry:

1. *How do you expect the demand (turnover) for <u>accommodation</u> of your company change over the next 3 months? It will…*
   + increase
   = remain unchanged
   − decrease

2. *How do you expect the demand (turnover) for <u>travel agency, tour operator reservation service and related activities</u> of your company change over the next 3 months? It will…*
   + increase
   = remain unchanged
   − decrease

The higher the tourism-specific expectations are, then this signifies that tourism-related businesess are more optimistic about their future demand (turnover) of their company. More details regarding the survey's questionnaire design, its reliability, the sample selection, and the processing of responses are provided by the European Union (2021). Our prior belief is that these two distinct survey-based expectations have predictive ability on future tourist arrivals.

From the above distinct forward-looking survey questions (showing expectations), we employ a Principal Component Analysis (PCA), and we construct a composite tourism-specific leading indicator (LEADING INDICATOR). In particular, first, we obtain the eigenvalue and eigenvector of their covariance matrix. We then construct the leading indicator index as a linear combination of the two variables by using the eigenvector associated with the largest eigenvalue as the corresponding weight. This approach has been widely used in the literature on the construction of sentiment and/or leading indicators (see among others, Chen *et al.,* 2014; Zhang *et al.,* 2019; Anastasiou and Katsafados, 2020; Anastasiou and Drakos, 2021; Anastasiou *et al.,* 2021). The first principal component derived from the PCA method explains 69.2% of the (standardized) sample variance, and only the first eigenvalue is far above 1.00, so we conclude that one factor captures the main variation.

Apart from the two above-mentioned main variables under scrutiny, we also incorporate in our model other factors that may affect tourist arrivals in Greece. First, as the purchasing power of people in one country positively affects their ability and inclination to travel to another country, we include a proxy for income. Given that the frequency of the dependent variable (inbound tourists) is on a monthly frequency, the use of quarterly or annual GDP may result in a significant loss of information regarding shorter-term variations. We thus opt to employ the monthly industrial production index (IPI) as a proxy for income (Nguyen and Valadkhani, 2020).

The cost of living at the destination relative to the origin is another critical factor that must be included. Given that tourists incur specific costs at the place of their destination, they compare prices between the destination and their home country; as a result, their decision on whether to visit a destination (Greece in our case) depends on the relative costs of living. The variable considered in this study and that has been used in the international tourism literature is the tourists' cost of living, defined as the annual percentage change of the Harmonized Consumer Price Index for the destination country (INFL) relative to the origin country (see among others, Song *et al.,* 2010; Gounopoulos *et al.,* 2012; Agiomirgianakis and Sfakianakis, 2014).

Following, among others, Chatziantoniou *et al.* (2016), Dragouni *et al.* (2016), Tsui *et al.* (2018), Nguyen and Valadkhani, (2020), we also consider the Greek Economic Policy Uncertainty (EPU) index of Hardouvelis *et al.* (2018) which can be found at the site of Baker *et al.* (2016). EPU is a proxy of the so-called crisis sentiment, and it captures concerns about the future state of the economy, thus reflecting changes in economic confidence.

The exchange rate is significant for foreign tourist inflows. To this end, following the standard practice in the literature (Lee and Chang, 2008; Lee *et al.,* 2021; Alola *et al.,* 2021; Mertzanis and Papastathopoulos, 2021), we also use the real effective exchange rate[5] (REER) relative to the effective exchange rate of Greece to capture the effect of relative prices, as well as country's international competitiveness.

[5] Coshall (2000) has found that raising the real effective exchange rate can also increase travel costs, bring in fewer tourists and be further harmful to tourism development.

Transportation costs may also well affect tourist arrivals. Given the fact that Greece is accessible by almost all alternative ways of traveling (at least from many European countries of origin), we opted for a more general proxy for the transportation cost, i.e., the international price of Brent crude oil (OIL) which would be expected to affect the costs of all of them (Agiomirgianakis and Sfakianakis, 2014).

Finally, note that given the monthly frequency of our dataset, all data used are seasonally adjusted. Series that were not available in this format from the data vendor, were transformed into seasonally adjusted series using the X12-Arima procedure provided by the US Census Bureau, which is a standard practice in the tourism literature (Cuccia and Rizzo, 2011). Besides, seasonality is one of the main aspects affecting tourism. According to Cuccia and Rizzo (2011), even if the seasonality of the tourism demand is trivial over time, the patterns of a given tourism destination's seasonality may have economic effects in terms of both social and private costs. Table 1 reports the definition and the main descriptive statistics of all the variables described above.

****Insert Table 1 here****

Figure 1 displays the time trajectory of each variable entered the VAR system. Figure 2 depicts the scatterplot between inbound international tourist arrivals and the tourism-specific leading indicator, from which a clear positive correlation is apparent. From both figures, we observe a common time path and a positive association between the two under-examination variables, providing tentative evidence confirming our priors.

****Insert Figures 1 & 2 here****

3.2 Methodology

Our analysis is based on a reduced-form VAR model, which is a system of equations where all variables are treated as endogenous, with the current values of the variables regressed against lagged values of all the variables in the system. Our specification contains seven variables[6] revolving around a core VAR model (see among others, Song and Witt, 2006; Gunter and Önder, 2016; Cao *et al.,* 2017). Before we embark on the estimation, as a first step of the empirical analysis, we examine all the variables for unit roots, performing the Augmented Dickey-Fuller (1979) test (ADF test), and the Phillips-Perron (1988) test (PP test). The empirical findings from the stationarity tests are reported in Table 2.

****Insert Table 2 here****

Then, we proceed with the choice of the appropriate lag length for the VAR specification, utilizing the Final Prediction Error (FPE), the Akaike Information Criterion (AIC), the Schwarz Bayesian Criterion (SBC), and the Hannan-Quinn Information Criterion (HQ) to determine the lag length of the VAR model. Table 3 shows that according to the Lag Order Selection Criteria, the appropriate lag length for the estimation of the VAR model is 2 lags[7].

****Insert Table 3 here****

Algebraically, the reduced-form finite-order VAR representation is as follows:

$$\boldsymbol{Y}_t = \boldsymbol{A_o} + \sum_{j=1}^{q=2} \boldsymbol{A}_j \boldsymbol{Y}_{t-j} + \varepsilon_t \, , \varepsilon_t \sim N(0, \boldsymbol{\Omega}) \tag{1}$$

where $\boldsymbol{Y}_t$ equals a (n×1) vector of variables under-scrutiny, $\mathbf{A_0}$ equals an (n×1) vector of constant terms, $\mathbf{A}_j$ denotes matrices of coefficients, **q** stands for the lag length, and ε_t denotes the

[6] All the variables used in the VAR model steam from the past empirical literature. It might be the case where some other additional variables should be incorporated in the system that may well affect tourism demand. However, in a VAR model, the number of parameters to be estimated grows exponentially with every additional endogenous variable, as the additional variable will result in an additional equation to be estimated. Hence, the estimation of a VAR model can become biased or unrealistic if the number of endogenous variables is large, explaining why other previous related studies also tend to deal with a relatively small number of variables (Song and Witt, 2006; Gunter and Önder, 2016; Cao *et al.,* 2017; Hamilton, 2020).

[7] According to the Schwarz Bayesian Criterion (SBC), one lag should be preferred. For a robustness check, we have also estimated an unrestricted VAR model with 1 lag instead of 2 lags, the results of the former remaining the same with those of the latter.

vector of residuals whose variance-covariance is $\boldsymbol{\Omega}$. The estimation method of the reduced-form VAR is OLS. Identification is achieved by Cholesky-decomposing the variance-covariance matrix of the VAR residuals, $\Omega = \mathbf{PP}'$, where $\mathbf{P}$ is the unique lower-triangular Cholesky factor with non-negative diagonal elements.

An important feature that a VAR model should meet is the so-called stability conditions. Figure 3 shows that the VAR model meets the stability conditions since the inverse roots of the AR characteristic polynomial lie inside the unit circle. Hence, we can infer that the VAR model is stationary and, thus, stable.

****Insert Figure 3 here****

4. Empirical Findings

Our results show that the one-period lag of the LEADING INDICATOR has the expected positive association with TOURIST and is statistically significant ($p < 0.01$), indicating that higher levels of tourism-specific business expectations in the previous month are associated with higher tourist arrivals in the future at all conventional significance levels. This finding is in line with the findings of Gholipour and Foroughi (2020), who also find a positive and statistically significant relationship between tourism demand and the general business confidence indicators. Hence, the proposed composite business indicators serves as a leading indicator and ultimately as a measure of tourism company managers' optimism or pessimism toward their near-future business performance.

We also find that the coefficient of the one-period lag of the dependent variable (TOURIST(–1)) is positive and statistically significant ($p < 0.01$), implying that current business travels are positively affected by the previous month's tourist arrivals and therefore there is evidence of persistence. In other words, an increase in tourist arrivals in the previous month will have a significant and prolonged impact on the future tourist arrivals' trajectory. The results also show that the model fits the data well according to the adjusted R-Squared values and the F-Statistics.

****Insert Table 4 here****

Unlike previous studies, we do not limit the analysis to the investigation of the business confidence as a leading indicator to explain the demand function of tourist arrivals, but we also generate shocks from this variable through the impulse response function by utilizing the VAR. To determine the statistical reliability of the response, Monte Carlo Simulation is used to construct the confidence bands around the impulse response. In order to determine the robustness and reliability of the response, we compute confidence bands using Monte Carlo Simulation that is simulated 200 times as a robustness test of the impulse response. The selection of the 200 simulations was based on the estimation sample and some recent empirical literature (see, among others, Galariotis *et al.,* 2016; Anastasiou and Drakos, 2021; Anastasiou *et al.,* 2021). This approach also allows evidence of a statistically significant response to the shock inflicted whenever the zero line lies outside the confidence bands.

Given that tourism demand is stationary, the impulse response should tend towards zero as the time period increases. Figure 4 demonstrates the the IRF derived from the VAR model, with the shocks measured by the Cholesky one standard deviation innovations. The results provide a clear picture of the impact of tourism-specific expectations' shocks on future tourism demand to Greece. Consistent with previous studies on the effect of business confidence on tourism demand, a one standard deviation of tourism expectations' shock originating from the leading indicator has an immediate positive impact on future tourists' arrivals. A closer inspection of the IRF reveals that the duration of the shock is highly persistent since, after its abrupt increase in the first four months, it continues to affect tourist arrivals for more than a year before it turns insignificant.

****Insert <u>Figure</u> <u>4</u> here****

The intuition behind our finding is that when managers of domestic tourism businesses form more optimistic expectations about their business performance in the near future, they may be more willing to search for more business opportunities in international markets, increasing their spending and ultimately attracting more international tourists. Subsequently, Greece's international tourist arrivals would increase, ceteris paribus. In other words, and in line with prior literature, business confidence due to increased uncertainty of future economic conditions or pessimism leads to an adverse reaction in future tourist arrivals. According to our results, this hypothesis is supported by the data. Furthermore, the implication of the impulse response results is to enhance the empirical finding that the predictive power of tourism demand models is improved by including the information contained in the tourism-specific leading indicator.

It is known that large confidence intervals around the impulse response call into doubt the reliability of the measurement information and the robustness of the response. The positioning of the confidence bands (Figure 4) suggests that the impulse responses are not very large and hence they are reliable and robust,. Therefore, the results of the impulse response analysis can be useful to practitioners since they show in a consistent manner how long and how intense the shocks are likely to have an impact.

To evaluate the contribution of each driver in the trajectory of Greek international tourist arrivals, we proceed to the estimation of the variance decomposition. Concerning the Forecast Error Variance Decompositions (FEVDs hereafter), the results are provided in Table 5, in which we report the variance of tourist arrivals (as a percentage) that is explained by each variable.

****Insert <u>Table</u> <u>5</u> here****

According to the FEVDs results, we find that tourism expectations explains a significant proportion of the variation of tourist arrivals. In particular, tourism expectations have, on average, 3 times greater proportion on explaining the variation of tourist arrivals compared to the majority of the rest of the variables.

4.1 Forecasting Exercise

To assess the predictive ability of the leading indicator in comparison with that of other econometric models, we perform a dynamic out of sample forecasting exercise for one-, two-, three-, and six- months ahead forecast horizons are considered. The measures of forecasting accuracy used are the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE)[8], the mathematical formulation of which reads as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \tag{2}$$

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{3}$$

where y_t and $\hat{y}_t$ denote the actual and the predicted value of tourist arrivals, respectively.

To compare the forecasting performance of the leading indicator, we examine three alternative models: (i) the VAR model, as described above; (ii) a VAR model where the leading indicator is excluded from the system, and hence tourist arrivals are explained only by the rest of the variables; and (iii) an ARIMA model. In an ARIMA (p,d,q) model, p, d, and q are integers greater than or equal to zero and refer to the order of the autoregressive, integrated, and moving average parts of the model, respectively. In our case, an ARIMA (4,1,1) is selected based on the results from the automatic ARIMA selection procedure we employed, based on the classical information criteria.

[8] The Mean Absolute Error (MAE) measures the degree to which forecasts, and the outcomes are close together, whereas the Root Mean Square Error (RMSE) is a measure of precision based on the residuals aggregated over the back-test period. Both measures of accuracy are frequently used in the literature (for instance, Athanasopoulos and Hyndman, 2008; Chu, 2009; Anastasiou and Drakos, 2021; Anastasiou *et al.,* 2021).

The results reported in Table 6 show each model's performance in terms of forecasting Greek tourist arrivals. We find that the proposed tourism-specific leading indicator not only serves as a leading indicator, explaining the variation of tourist arrivals, but also reduces both RMSE and MAE. The VAR model incorporating the leading indicator outperforms all competing models across every forecasting horizon. On the other hand, the worst performing model based on these criteria is the ARIMA (4,1,1) model, a finding that is robust as the forecasting period expands. This finding is consistent with the results of Smeral and Wuger (2005) and Gounopoulos *et al.* (2012) who reported that other models outperformed the ARIMA model.

****Insert Table 6 here****

5. Concluding Remarks

We construct a leading indicator based on business expectations pertinent to the tourism sector serving as a measure of tourism businesses' optimism or pessimism toward their near-future business performance. Using monthly data over the period of 2002-2021 and a VAR model, we find that the proposed leading indicator exhibits a strong predictive power for future tourist inflows in the country. Furthermore, our findings from the Forecast Error Variance Decomposition also suggest that the tourism-specific expectations explain a significant proportion of the variation of tourist arrivals over time.

Apart from investigating the impact of tourism-specific business expectations on tourist arrivals, we also estimate impulse response functions within the VAR framework. We find that a positive shock by one standard deviation in the leading indicator leads to a persistent positive response of future tourist arrivals. Therefore, policymakers should monitor this leading indicator to better capture the dynamics of tourist flows. In addition, our results highlight the strong potential of this new leading indicator to improve the forecasting power in the case of tourism.

Accurate forecasts can offer valuable support to businesses in making the most strategic decisions during peak or off-peak tourism seasons.

Overall, the results of the present study can benefit government officials, forecasters and policymakers in a variety of ways. For forecasters, it is a guide to obtain the best out-of-sample forecasts in a multivariate framework. For government officials and policymakers, the usage of a leading indicator making accurate predictions of international tourist arrivals can promote the planning of optimum policies, resource allocations and investment decisions related to tourism.

References

Agiomirgianakis, G. M., & Sfakianakis, G. (2014). Determinants of tourism demand in Greece: a panel data approach. *Ekonometria*, *1*, 15-26.

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: capitalizing on big data. *Journal of Travel Research*, 58(2), 175-191.

Alola, A. A., Uzuner, G., & Akadiri, S. S. (2021). Modeling tourism and fear nexus in G4 countries. *Current Issues in Tourism*, 24(10), 1333-1339.

Anastasiou, D., & Drakos, K. (2021). European depositors' behavior and crisis sentiment. *Journal of Economic Behavior & Organization*, 184, 117-136.

Anastasiou, D., & Katsafados, A. G. (2020). Bank Deposits Flows and Textual Sentiment: When an ECB President's speech is not just a speech. MPRA Working Paper, No. 99729.

Anastasiou, D., Kallandranis, C., & Drakos, K. (2022a). Borrower discouragement prevalence for Eurozone SMEs: Investigating the impact of economic sentiment. *Journal of Economic Behavior & Organization*, 194, 161-171.

Anastasiou, D., Ballis, A., & Drakos, K. (2022b). Constructing a positive sentiment index for COVID-19: Evidence from G20 stock markets. *International Review of Financial Analysis*, 81, 102111.

Anastasiou, D., Kapopoulos, P., & Zekente, K. M. (2021). Sentimental Shocks and House Prices. *The Journal of Real Estate Finance and Economics*, 1-29.

Athanasopoulos, G., & Hyndman, R.J. (2008). Modelling and forecasting Australian domestic tourism, *Tourism Management,* 29, 19-31.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4), 1645-1680.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593-1636.

Cao, Z., Li, G., & Song, H. (2017). Modelling the interdependence of tourism demand: The global vector autoregressive approach. *Annals of Tourism Research*, 67, 1-13.

Chatziantoniou, I., Degiannakis, S., Eeckels, B., & Filis, G. (2016). Forecasting tourist arrivals using origin country macroeconomics. *Applied Economics*, 48(27), 2571-2585.

Chen, H., Chong, T. T. L., & She, Y. (2014). A principal component approach to measuring investor sentiment in China. *Quantitative Finance*, 14(4), 573-579.

Chen, N., Hsu, C. H., & Li, X. (2021). Resident sentiment toward a dominant tourist market: scale development and validation. *Journal of Travel Research*, 60(7), 1408-1425.

Choi, J. G. (2003). Developing an economic indicator system (a forecasting technique) for the hotel industry. *International Journal of Hospitality Management*, 22(2), 147-159.

Chu, F. L. (2009). Forecasting tourism demand with ARMA-based methods. *Tourism Management*, 30(5), 740-751.

Coshall, J. T. (2000). Spectral analysis of overseas tourists' expenditures in the United Kingdom. *Journal of Travel Research*, 38, 292–298.

Crotts, J. C., Thunberg, E. M., & Shifflet, D. K. (1993). Consumer Confidence as a Leading Indicator of Change in US Travel Volume. *Journal of Travel & Tourism Marketing*, 1(2), 53-62.

Cuccia, T., & Rizzo, I. (2011). Tourism seasonality in cultural destinations: Empirical evidence from Sicily. *Tourism Management*, 32(3), 589-595.

Darling, P. G. (1955). A Surrogative Measure of Business Confidence and Its Relation to Stock Prices. *Journal of Finance,* 10, 442–58.

Dragouni, M., Filis, G., Gavriilidis, K., & Santamaria, D. (2016). Sentiment, mood and outbound tourism demand. *Annals of Tourism Research*, 60, 80-96.

European Union. (2021). The Joint Harmonised EU Programme of Business and Consumer Surveys. User Guide, October 2021.

Fu, Y., Hao, J. X., Li, X., & Hsu, C. H. (2019). Predictive accuracy of sentiment analytics for tourism: A metalearning perspective on Chinese travel news. *Journal of Travel Research*, 58(4), 666-679.

Galani-Moutafi, V. (2004). Tourism research on Greece. *Annals of Tourism Research,* 31, 157-179.

Galariotis, E. C., Makrichoriti, P., & Spyrou, S. (2016). Sovereign CDS spread determinants and spill-over effects during financial crisis: A panel VAR approach. *Journal of Financial Stability*, 26, 62-77.

Gholipour, H. F., & Foroughi, B. (2020). Business sentiment and international business travels: a cross-country analysis. *Journal of Travel Research*, 59(6), 1061-1072.

Gounopoulos, D., Petmezas, D., & Santamaria, D. (2012). Forecasting tourist arrivals in Greece and the impact of macroeconomic shocks from the countries of tourists' origin. *Annals of Tourism Research*, 39(2), 641-666.

Guizzardi, A., & Stacchini, A. (2015). Real-time forecasting regional tourism with business sentiment surveys. *Tourism Management*, 47, 213-223.

Gunter, U., & Önder, I. (2016). Forecasting city arrivals with Google Analytics. *Annals of Tourism Research*, 61, 199-212.

Hamilton, J. D. (2020). Time series analysis. Princeton university press.

Hao, J. X., Fu, Y., Hsu, C., Li, X., & Chen, N. (2020). Introducing news media sentiment analytics to residents' attitudes research. *Journal of Travel Research*, 59(8), 1353-1369.

Hardouvelis, G. A., Karalas, G., Karanastasis, D., & Samartzis, P. (2018). Economic policy uncertainty, political uncertainty and the Greek economic crisis. Political Uncertainty and the Greek Economic Crisis. Available at SSRN: https://ssrn.com/abstract=3155172

Hassani, H., Silva, E. S., Antonakakis, N., Filis, G., & Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research*, 63, 112-127.

Icoz, O., Kozak, M. & Var, T. (1998). Tourism demand in Turkey, *Annals of Tourism Research*, 25, 236-240.

Adamopoulou E., Kapopoulos, P., & Marinopoulou E., (2022). Greek Tourism Industry Reloaded: Post-pandemic rebound, travel megatrends and future challenges, *Insights*, Alpha Bank.

Khan, H., & Upadhayaya, S. (2020). Does business confidence matter for investment?. *Empirical Economics*, 59(4), 1633-s1665.

Kim, H. B., Park, J. H., Lee, S. K., & Jang, S. S. (2012). Do expectations of future wealth increase outbound tourism? Evidence from Korea. *Tourism Management*, 33(5), 1141-1147.

Kim, J. H., & Moosa, I. A. (2005). Forecasting international tourist flows to Australia: a comparison between the direct and indirect methods. *Tourism Management*, 26, 69-78.

Kulendran, N., & Witt, S. F. (2003). Leading indicator tourism forecasts. *Tourism Management*, 24(5), 503-510.

Lee, C. C., & Chang, C. P. (2008). Tourism development and economic growth: A closer look at panels. *Tourism Management*, 29(1), 180-192.

Lee, C. C., Olasehinde-Williams, G., & Akadiri, S. S. (2021). Geopolitical risk and tourism: Evidence from dynamic heterogeneous panel models. *International Journal of Tourism Research*, 23(1), 26-38.

Letdin, M., Sirmans, S., & Sirmans, G. S. (2021). Betting Against the Sentiment in REIT NAV Premiums. *The Journal of Real Estate Finance and Economics*, 1-25.

Lim, C., Chang, C., & McAleer, M. (2009). Forecasting h(m)otel guest nights in New Zealand. *International Journal of Hospitality Management*, 28(2), 228-235.

Lolos, S., Palaios, P., & Papapetrou, E. (2021). Tourism-led growth asymmetries in Greece: evidence from quantile regression analysis. *Portuguese Economic Journal*, 1-24.

Mertzanis, C., & Papastathopoulos, A. (2021). Epidemiological susceptibility risk and tourist flows around the world. *Annals of Tourism Research*, 86, 103095.

Nguyen, J., & Valadkhani, A. (2020). Dynamic responses of tourist arrivals in Australia to currency fluctuations. *Journal of Hospitality and Tourism Management*, 45, 71-78.

Papatheodorou, A., & Song, H. (2005). International tourism forecasts: A time series analysis of world and regional data, *Tourism Economics,* 11, 11-24.

Phillips, P.C.B & Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, 75, 335-346.

Saayman, A., & Saayman, M. (2008). Determinants of inbound tourism to South Africa. *Tourism Economics*, 14(1), 81-96.

Santamaria, D., & Filis, G. (2019). Tourism demand and economic growth in Spain: New insights based on the yield curve. *Tourism Management*, 75, 447-459.

Smeral, E., & Weber, A. (2000). Forecasting international tourism trends to 2010, *Annals of Tourism Research*, 27, 982-1006.

Smeral, E., & Wuger, M. (2005). Does complexity matter? Methods for improving forecasting accuracy in tourism: The case of Australia. *Journal of Travel Research*, 44, 100-110.

Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tourism Management*, 27(2), 214-224.

Song, H., Li, G., Witt, S.F. & Athanasopoulos, G. (2010). Forecasting tourist arrivals using time-varying parameter structural time series models, *International Journal of Forecasting,* 27(3), 855-869.

Song, H., Wong, K. K., & Chon, K. K. (2003). Modelling and forecasting the demand for Hong Kong tourism. *International Journal of Hospitality Management*, 22(4), 435-451.

Tavares, J. M., & Leitao, N. C. (2017). The determinants of international tourism demand for Brazil. *Tourism Economics*, 23(4), 834-845.

Taylor, K., and McNabb, R. (2007). Business Cycles and the Role of Confidence: Evidence for Europe. *Oxford Bulletin of Economics and Statistics,* 69, 185–208.

Tsui, W. H. K., Balli, F., Tan, D. T. W., Lau, O., & Hasan, M. (2018). New Zealand business tourism: exploring the impact of economic policy uncertainties. *Tourism Economics*, 24(4), 386-417.

Yost, E., Ridderstaat, J., & Kizildag, M. (2020). Early warning indicators? The effect of consumer and investor sentiments on the restaurant industry. *International Journal of Hospitality Management*, 89, 102575.

Zhang, R., Xian, X., & Fang, H. (2019). The early-warning system of stock market crises with investor sentiment: Evidence from China. *International Journal of Finance & Economics*, 24(1), 361-369.

Tables

Table 1: Variables' definition and descriptive statistics of data

Variable	Definition	Proxy	Source	Mean	Median	Minimum	Maximum	Std. Dev.
TOURIST	International (foreign country) tourist arrivals, incl. Hotels; holiday and other short-stay accommodation; camping grounds, recreational vehicle parks and trailer parks (in millions of people)	International tourist arrivals	EUROSTAT	1.018	0.585	0.004	4.896	1.087
LEADING INDICATOR	The common factor of the businesses' expectations of the demand over the next 3 months for (i) accommodation; and (ii) travel agency, tour operator reservation service and related activities	Tourism expectations	EUROPEAN COMMISSION	0.001	0.215	-4.891	3.015	1.164
REER	Real Effective Exchange Rate	Competitiveness	EUROSTAT	95.610	95.632	86.131	102.784	3.289
OIL	Brent crude oil price	Transportation cost	Federal Reserve Bank of St. Louis	67.342	63.428	18.378	132.718	28.123
INFL	Annual growth rate of the Harmonized Index of Consumer Prices	Cost of living	EUROSTAT	1.735	1.915	-0.595	4.478	1.028
IPI	Industrial Production Index	Income	EUROSTAT	119.579	111.050	149.800	95.300	17.761
EPU	Economic Policy Uncertainty of Hardouvelis *et al.* (2018)	Economic Uncertainty	Website of Baker *et al.* (2016) https://www.policyuncertainty.com	98.991	94.530	188.700	37.700	28.121

Notes: This table presents the definition and the descriptive statistics of each variable used in the analysis. Descriptive statistics stand for the original (i.e., non-transformed) data.

Table 2: Unit Root tests

	Augmented Dickey-Fuller test statistics		Phillips-Perron test statistic	
Variable	Original series	Transformed series	Original series	Transformed series
TOURIST	-2.617*	-5.000***	-2.334	-5.079***
LEADING INDICATOR	-4.835***	-	-4.694***	-
REER	-1.533	-4.244***	0.126	-3.126**
OIL	-4.960***	-3.132**	-4.286***	-2.832*
INFL	-1.860	-	-1.860	-
IPI	-1.417	-25.496***	-1.435	-27.231 ***
EPU	-4.557***	-4.189***	-6.898***	-6.149***

Notes: The null hypothesis in each test is that the variable is unit root. The hypothesis is accepted in the log levels of some variables but rejected (as expected) after taking the first differences (Δlog). Therefore, we uncover stationarity in the transformed series.

Table 3: VAR Lag Order Selection Criteria

Lag	FPE	AIC	SBC	HQ
0	0.000686	12.58019	12.69102	12.62499
1	8.88e-09	1.325104	**2.211750***	1.683466
2	**5.14e-09***	**0.777396***	2.439856	**1.449324***
3	6.29e-09	0.976231	3.414506	1.961725
4	7.08e-09	1.087212	4.301302	2.386272
5	7.62e-09	1.149981	5.139885	2.762607
6	8.49e-09	1.240629	6.006348	3.166822
7	6.94e-09	1.016401	6.557935	3.256159
8	8.58e-09	1.196872	7.514221	3.750197
9	8.26e-09	1.117259	8.210422	3.984150
10	9.06e-09	1.157329	9.026308	4.337787
11	9.33e-09	1.121685	9.766478	4.615709
12	9.74e-09	1.084805	10.50541	4.892394

Notes: FPE, AIC, SBC, and HQ stand for the Final prediction error, the Akaike information criterion, the Schwarz information criterion, and the Hannan-Quinn information criterion, respectively. Three out of the four information criteria suggest the selection of 2 lags.

Table 4: Estimation results from the VAR model

	TOURIST
TOURIST(-1)	0.865*** (0.065)
TOURIST(-2)	-0.174*** (0.069)
LEADING INDICATOR(-1)	0.099*** (0.039)
LEADING INDICATOR(-2)	-0.044 (0.038)
Other endogenous variables	Included
Diagnostics	
Number of Observations	222
R2-adjusted	73.5%
F-Statistic	44.679***
Serial Correlation LM Test (p_value)	0.175

Notes: This Table shows the estimation results from the VAR model with 2 lags. For brevity, we report only the under examination variables. The asterisks *** imply statistically significant coefficients at the 1% level of significance.

Table 5: Forecast-error variance decompositions for TOURIST

Forecast Period	TOURIST	LEADING INDICATOR	EPU	IPI	REER	OIL	INFL
1	100.000	0.000	0.000	0.000	0.000	0.000	0.000
2	90.921	1.706	0.083	0.056	0.846	6.298	0.092
3	82.455	3.453	0.127	0.289	1.133	12.472	0.072
4	78.136	5.000	0.252	0.643	1.035	14.841	0.093
5	76.265	5.931	0.451	1.011	1.226	15.008	0.109
6	75.092	6.372	0.678	1.322	1.698	14.729	0.109
7	74.236	6.552	0.881	1.575	2.148	14.502	0.107
8	73.637	6.624	1.040	1.791	2.461	14.341	0.106
9	73.207	6.650	1.160	1.986	2.663	14.229	0.105
10	72.876	6.652	1.250	2.164	2.801	14.150	0.106
11	72.609	6.643	1.320	2.328	2.901	14.092	0.107
12	72.388	6.628	1.374	2.479	2.976	14.046	0.109
Notes: This table shows the forecast-error variance decompositions for the variable TOURIST.							

Table 6: Forecasting Tourist Arrivals – Model Performance

Period	Model	RMSE	MAE
1-month ahead	VAR incl. the leading indicator	**0.16722**	**0.15684**
	VAR excl. the leading indicator	0.31642	0.29720
	ARIMA (4,1,1)	0.23962	0.22864
2-months ahead	VAR incl. the leading indicator	**0.44899**	**0.42789**
	VAR excl. the leading indicator	0.55357	0.52622
	ARIMA (4,1,1)	0.74839	0.72602
3-months ahead	VAR incl. the leading indicator	**0.54871**	**0.49561**
	VAR excl. the leading indicator	0.60208	0.54305
	ARIMA (4,1,1)	0.74604	0.65025
6-months ahead	VAR incl. the leading indicator	**0.38221**	**0.34840**
	VAR excl. the leading indicator	0.43129	0.38271
	ARIMA (4,1,1)	0.50504	0.40675

Notes: This table shows each model's performance in terms of forecasting tourist arrivals in Greece. The measures of forecasting accuracy used are the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

Figures

Figure 1: Time trajectory of each variable

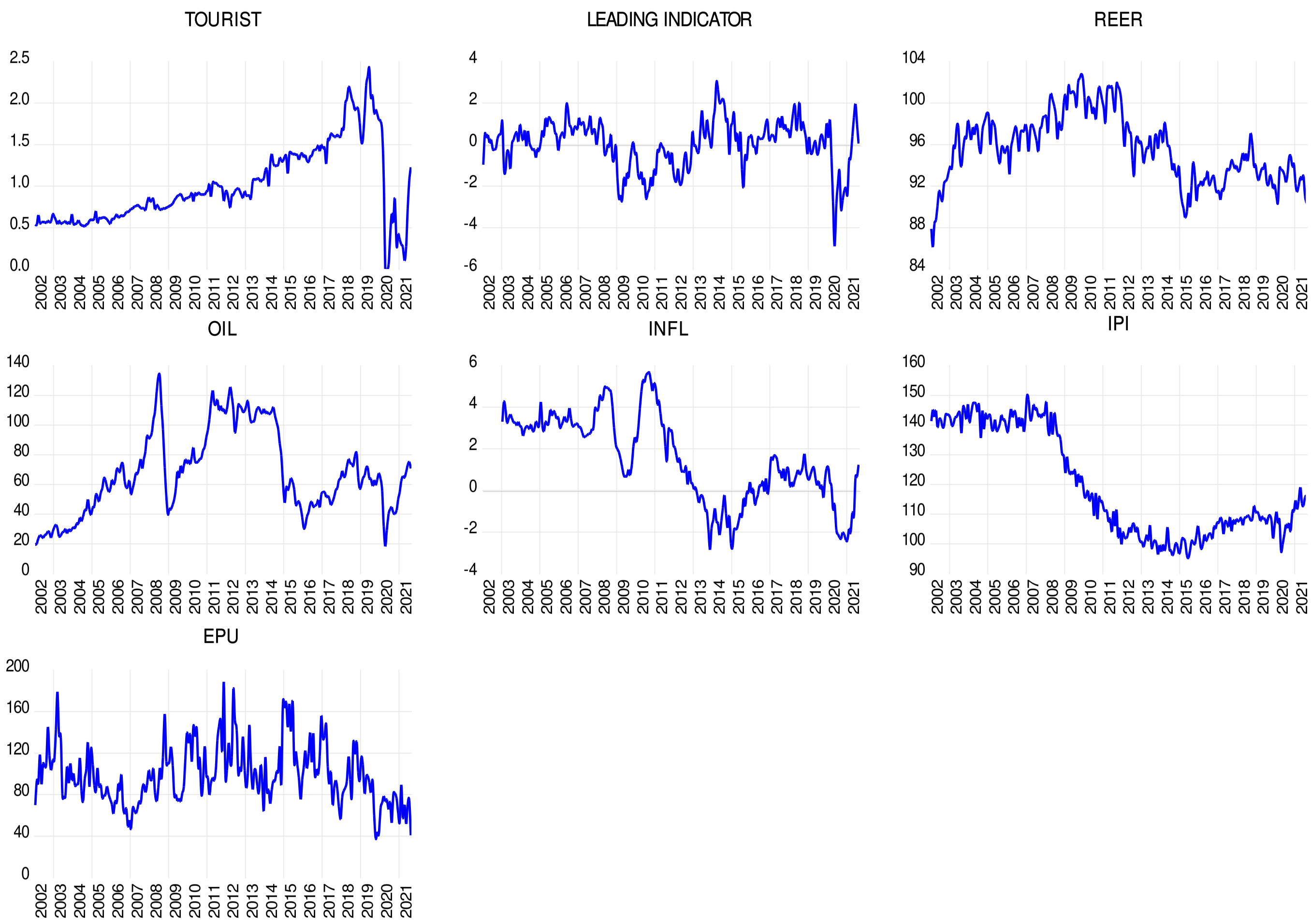

Figure 2: Scatterplot between inbound international tourist arrivals and the tourism-specific sentiment as leading indicator

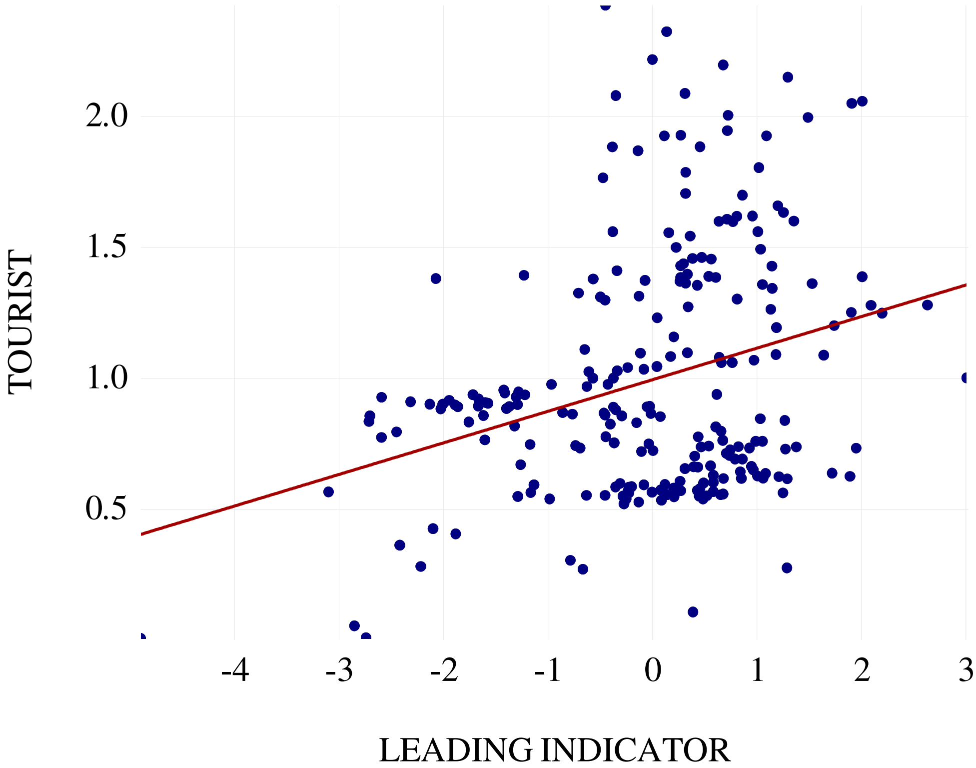

Figure 3: VAR Stability Conditions

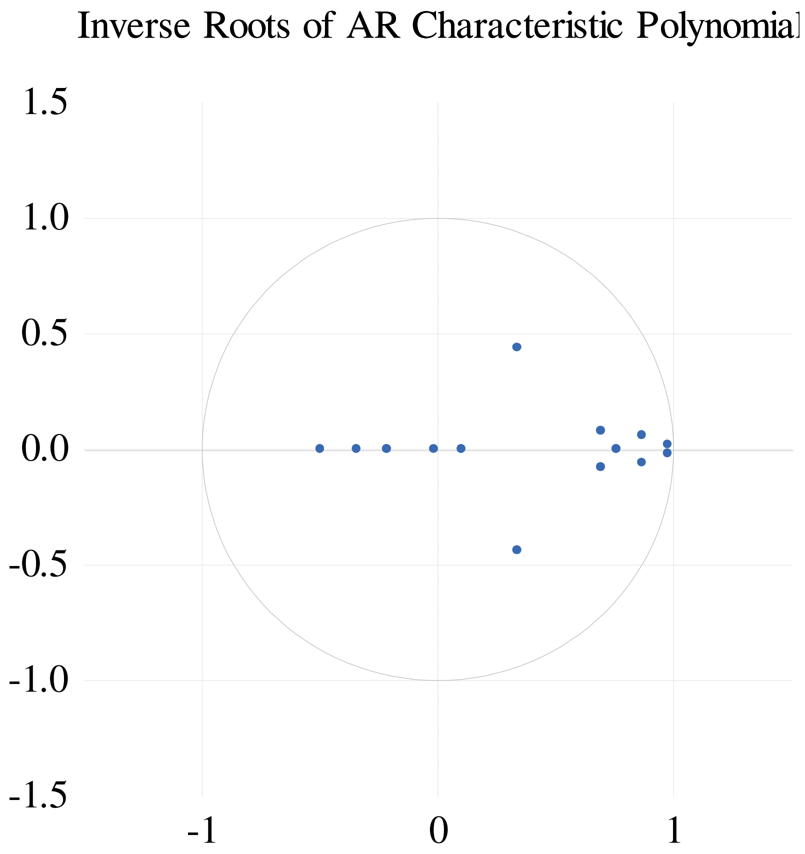

Figure 4: Impulse Response Functions

Response to Cholesky One S.D. (d.f. adjusted) Innovations ± 2 S.E.

Response of TOURIST to LEADING INDICATOR