ITJ | The International Tax Journal

Determinants of tourism demand in Tipaza (Algeria)

Mokrani Tayeb [1], **Dr. Zahali Rachid** [2] and **Dr. Touahria Walid** [3]

[1] Faculty of Economic Sciences and Management, University of Sfax, Tunisia
Email: tayebmokrani82@gmail.com
[2] Specialization: Service Marketing, Faculty of Economic, Commercial, and Management Sciences, University of Algiers 3, (Algeria). Email: zahalirachid18@gmail.com
[3] Institute of Economic, Commercial, and Management Sciences, Mersli Abdullah University Center, Tipaza, (Algeria). Email: waliding80@gmail.com

Abstract---This study aims, within the framework of promoting tourism in the Tipaza region, to shed light on tourism demand in the area in order to understand tourist behaviour and identify the key determinants. This is particularly relevant as Algeria was affected by the COVID-19 pandemic, which partially disrupted domestic tourism activities. These events lead researcher to consider the effectiveness of statistical methods in forecasting the recovery of tourism demand during an unexpected and volatile crisis, prompting the use of dynamic models. The study relies on a data set of monthly variables, including the number of tourists in the region (VT), the number of hotel nights (NT), the number of beds (NLT), the number of rooms (NCH), and employment (EMP) from Q1 2018 to Q4 2022, collected from the tourism sector in Tipaza. The findings indicate that seasonal employment positively contributes to tourism activity, whereas permanent employment has a negative impact. Additionally, the number of rooms plays a crucial role in determining hotel demand, as it is linked to the limited accommodation capacity during peak seasons. The study also reveals that tourism in Tipaza relies primarily on local demand, with public sector facilities concentrated in the eastern coastal strip, while the private sector dominates the rest of the coastline.

Keywords---tourism demand, ARDL, Tipaza.

Introduction:

At the academic level, the desire to use economic measurement techniques has become increasingly urgent and desirable in various fields of scientific research, especially in tourism demand studies. This is essential for understanding tourist behavior and identifying the determinants of tourism demand. Our

How to Cite:
Tayeb, M., Rachid, Z., & Walid, T. (2025). Determinants of tourism demand in Tipaza (Algeria). *The International Tax Journal*, *52*(3), 439–457. Retrieved from https://internationaltaxjournal.online/index.php/itj/article/view/72

The International tax journal ISSN: 0097-7314 E-ISSN: 3066-2370 © 2025

Submitted: 26 Nov 2024| Revised: 10 Jan 2025| Accepted: 05 May 2025

study focuses on the case of the Tourist region, Tipaza, which serves as a model of a tourist pole that is growing over time. This is particularly significant following the impact of the COVID-19 pandemic, which partially disrupted domestic tourism activities in Algeria.

These events lead the researcher to consider the effectiveness of statistical methods for predicting the recovery of tourism demand during sudden and volatile crises. Consequently, we have explored appropriate dynamic models for this purpose. In the context of tourism in the Tipaza region, we aim to identify the most important determinants that contribute to its selection as a domestic destination. To achieve this, we tested a sample data consisting of the following variables: the number of tourists to the state (VT), the number of hotel nights (NT), the number of beds (NLT), the number of rooms (NCH), and the workforce (EMP).

We also want to compare the number of tourists and the number of hotel nights as dependent variables, while treating the remaining variables as independent. Econometric analysis relies on several requirements and assumptions linked to modeling and estimation. This necessitates studying the dataset from multiple perspectives, beginning with a descriptive statistical analysis to assess the homogeneity of the data in terms of magnitude. It is also important to determine whether the time series experience any events, shocks, or structural changes that must be considered during modeling.

We will begin by verifying the condition of normal distribution of the variables, in addition to conducting graphic analysis to reveal the components of the time series, mainly the general trend, seasonality, and volatility. All of these factors are potential causes of time series instability, which must be addressed carefully.

Variables of Study

In order to examine tourism demand in Algeria, specifically in Tipaza Province as a model, we utilize a sample of variables that we assume as demand determinants. The variables related to tourism, investment structures, and employment are viewed as key contributors to local tourism demand. We gathered monthly data on visitors, hotel nights, beds, rooms, and workers from January 2018 to December 2022.

The study sample included five variables and sixty (60) observations, which seem heterogeneous in terms of units. Notably, according to the Jarque-Bera test, all study variables follow a normal distribution, aligning with the estimation method. However, the heterogeneity of the units, particularly between the number of workers and the number of nights, permits the transformation of the variables into their logarithmic form. This conversion is important, in terms of expression (elasticities) and interpretation.

Table (1): Model variables

Variable acronyme	Variable name	Nature
LNT	Number of nights	Dependent
LNLT	Number of beds	Independent
LNCH	Number of rooms	Independent
LEMP	Number of workers	Independent
LVT	Number of visitors	Dependent

It appears that these time series have undergone critical periods causing significant changes, particularly in 2020 marked by the COVID-19 pandemic crisis that drove widespread panic and extreme fear among various segments of society. The quarantine policies led to economic stagnation, during which numerous economic and service activities halted, and tourism sector globally affected; Algeria was no exception.

These series are characterized by random behavior, driven by general trend and a modest seasonality effect. The tourism sector faced tremendous shock due to the COVID-19 epidemy, driving tourism demand to decline to its lowest level in 2020. Subsequently, thanks to improvements in prevention and treatment, tourism demand began to recover, experiencing modest increases. The general trend component shifted from a negative to a positive tendency during the period of 2022-2023.

Calculating the total and partial autocorrelation functions reveal a clear seasonal effect with 12 periods, aligning with monthly data. The exception among these variables pertains to the operation, whose periodicity was not consistent with that of the other variables. This finding aids in selecting an appropriate unit root test that accounts for seasonality as a factor of instability, in addition to the global trend. We also exclude variance as a potential cause due to the logarithmic transformation.

Initial attempt to modelling and estimation

In most cases, when we want to design an economic model, we rely on economic theory if it is available and there is a consensus regarding the phenomenon to be studied. In our current situation, however, and in the absence of any relevant theory, we utilized a logical approach that we arranged in the form of the model shown below. To determine the functional form, we plotted the dependent variable against each independent variable individually, which led us to the conclusion that a linear form could be appropriate.

It is evident from these various plots that we can accept the linear form shared by these data. Therefore, before delving into the practical aspect and to avoid the potential issue of multicollinearity, we compared the number of tourists and the number of nights as a dependent variable, with the remaining variables serving as independent variables. After employing the Akaike Information Criterion (AIC) test, used to compare two models from the same formation (or Nested Model) with an equal number of parameters, we adopted the number of nights as the dependent variable expressing the number of tourists, as shown in the following model:

$$LNT = a_1 + a_2LNLT + a_3LNCH + a_4LEMP + ȩ$$

By using the study data and the least squares method, the following results were achieved:

$$LNT = \underset{(0.18)}{4.77} - \underset{(0.09)}{0.3LNLT} + \underset{(0.03)}{1.05LNCH} + \underset{(0.03)}{0.24LEMP}$$

$R2=0.93$ $Dw=0.17$ $T=60$

For decades, interest has focused on using econometric-based forecasting models to explore the causal relationships between economic factors and tourism demand. Econometric models differ from time series models in their ability to capture the causal relationships between tourism demand and various related factors, particularly those associated with tourism supply. This enables the explanation, analysis, and prediction of future tourism demand movements while attempting to identify the reasons for changes in demand, thereby allowing for informed tourism decision-making.

Despite the significance of this static estimate, the unexpected indication of the variable number of beds raised concerns. Specifically, it suggested that increasing the number of beds does not contribute to an increase in the number of tourists and may actually decrease it. This could be better understood in conjunction with the variable representing the number of rooms. We also worry about the possibility of correlation or multicollinearity, given the expected relationship between the number of beds and the number of rooms.

To gain a clearer understanding of the importance of this model in explaining tourism demand, it is necessary to support the above findings with additional statistical tests. To determine the statistical significance of this model, we initially employed the student's t-test to assess the significance of each parameter separately in the multiple linear model above. We utilized the well-known t-statistic, calculated as follows:

Table (2): Model estimation

Variable parameter	Estimated	Standard deviation	t_ Statistics	cv
Fixed limit a1	4.77	0.18	25.37	1.96
LNLT a2	-0.3	0.09	-3.25	1.96
LNCH a3	1.05	0.03	7.99	1.96
LEMP a4	0.24	0.03	7.82	1.96

Source: Eviews output.

Statistically, all the model's parameters were significant at the 5 percent level. However, economically, the number of rooms variable was negative, indicating that an increase in the number of beds leads to a decrease in the number of tourists due to overcrowding and declining service quality. This may also suggest the possibility of multicollinearity. We tend to exclude the number of beds variable and focus instead on the number of rooms as the most significant predictor of tourism demand, which led to its exclusion from the model. We obtained the following new estimates:

$LNT = 4.98 + 0.65LNCH + 0.25LEMP$
(0.19) (0.05) (0.03) R2=0.91 Dw=0.21 T=60

We now observe an improvement in the model in terms of sign and magnitude, with highly significant parameters. However, characteristics of a spurious regression is evident from the rule of thumb test, where DW > R^2. This diagnosis encompasses several explanations and potential reasons, with the most significant being the loss of dynamic effects in the model. Therefore, as a preliminary step, we will move towards examining the stability of the time series to address the issue of false estimates.

First: Testing the Stability of Time Series

Currently, time series primarily suffer from the issue of instability, also referred to as non- stationarity. This issue is particularly observed in time series that are influenced by general trends. Technically speaking, these series contain a unit root, meaning that the mean and variance of the variables are not independent of time. Therefore, it can be said that stable time series are characterized by a constant mean and variance over time, and the variance between two values in the series is only related to the length of time separating them.

There are two other reasons for instability: volatility in the time series, expressed as non- constant variance, meaning the variance becomes dependent on time (for example, it may increase as time progresses from one period to another), and seasonal effects, which refer to regular seasonal influences on the time series. These effects are clearly visible when plotting the autocorrelation and partial autocorrelation functions.

This phenomenon has encouraged the search for methods to test the stability of time series. However, in some cases, a superficial understanding of the stationarity of a time series can be achieved if its data fluctuates around a constant mean that is independent of time. Generally, to determine whether a time series is stable or not, we resort to a set of tests known in the literature as Unit Root Tests.

In practice, it can sometimes be difficult to determine the nature of a time series in terms of its components and stationarity, whether through simple observation, graphical representation, or even the autocorrelation function. Before the widespread use of unit root tests, the issue of instability was not as prominent as it is today. At that time, researchers relied on simple and straightforward methods, such as using basic statistical measures to test for the presence or absence of a general trend in the series. This was done using non-parametric tests like the sign test, runs test, and the Daniel test. Another approach involved checking the stability of the series' mean by dividing the time series into two equal parts and calculating the mean for each subset. If the means were equal, the series was considered stable, implying the absence of a general trend. However, if a significant difference was observed, the series was deemed unstable. The same test could also be applied using parametric methods, such as estimating the general trend through ordinary least squares regression.

Unit Root Test:
This test originates from the use of the autocorrelation function to determine the nature of instability in time series within the framework of Box-Jenkins models. It aims to select the most appropriate method to address instability, which is the direct cause of spurious regression. This can be done either by removing the trend component (detrending) or by differencing the series. The latter approach helps avoid spurious regression and preserves the properties of statistical tests, as unstable time series distort the characteristics of traditional tests and cause them to lose their usual distributions. Several widely used tests, often included in statistical software packages, are as follows:

A. The Simple Dickey-Fuller Test (DF):
This is one of the earliest tests used in this field, considered the simplest and most widely used in standard econometric research and forecasting techniques. It was developed by Dickey and Fuller in 1979 and is denoted as the Dickey-Fuller (DF) test. The test relies on three simple equations that assume a random context following an autoregressive pattern.

The test examines the null hypothesis of non-stationarity for the time series of the variable under study. If the absolute calculated t-value of the parameter is less than the critical value, the null hypothesis of non-stationarity is accepted, indicating that the series contains a unit root and is unstable at its original level. This is expressed as the series not being I(0), as commonly referred to in Box-Jenkins models. However, if the t-value is greater than the critical value, the time series is considered stable at its first or second difference (which requires confirmation by testing the series at first differences).

When using statistical software alike EViews, it is preferable to refer to the three equations. After an initial examination of the time series, the researcher can directly proceed to the last equation if they believe the series contains a trend component and a constant term. The significance of the two parameters is then tested to determine whether the series is stable or not. The first equation is typically revisited after the initial examination, especially if the series is purely random, lacks a trend component, and does not include a constant term. To avoid the pitfalls of traditional testing tools, which lose credibility due to the risk of spurious regression, new critical values calculated using the Monte Carlo simulation method are now included in most statistical software packages.

B. Augmented Dickey-Fuller Test (ADF):
In the quest for stability and stationarity, the simple Dickey-Fuller test faced some limitations, such as the need to exclude seasonal effects, autocorrelation, and structural changes in the time series. This led Dickey and Fuller to develop an enhanced version in 1981, known as the Augmented Dickey-Fuller (ADF) test. This version is designed to address the issue of autocorrelation by incorporating three equations: one with a constant and trend, one with only a constant, and one with neither a constant nor a trend. The latter is used in cases where cointegration is being tested.
The results of this test reveal whether the time series is stable, as well as the nature of the trend component—whether it is systematic or random. Each case requires a specific approach during modelling or when transforming the series into a stationary state to avoid spurious regression.

C. Phillips-Perron Test:

To address the issue of heteroskedasticity in errors, which the previous two tests could not handle, Phillips and Perron proposed this test in 1988. The Phillips-Perron test corrects for heteroskedasticity in the residuals of the unit root test equation using a non-parametric method. It takes into account the conditional variance of errors, allowing it to eliminate biases caused by random fluctuations. It is also used to verify the integration of time series. This test follows the same statistical distribution as the Dickey-Fuller test, so the same critical values are used for both. However, it differs from the previous tests in that the null hypothesis indicates stationarity, unlike the Dickey-Fuller test, where the null hypothesis indicates non-stationarity.

KPSS Test (1992):

This test is based on the well-known Lagrange Multiplier (LM) test, which has the null hypothesis of stationarity and the alternative hypothesis of non-stationarity. The null hypothesis is rejected if the LM test statistic is greater than the critical value, i.e., LM > KPSS_cv. Using software like EViews simplifies these tests, as they are presented in a standardized format. It is important to note that a decision on stationarity can only be made if the estimated model has statistically significant variables.

Second: Cointegration

The concept of cointegration was introduced in 1987 by Engle and Granger. It is crucial in studying the statistical properties of time series, particularly regarding long-term equilibrium. The cointegration model assumes a balanced relationship between economic variables in the long run, where any short-term deviation from equilibrium is corrected by economic forces that bring the variables back to long-term equilibrium. Among the most important tests for cointegration in time series are the Engle-Granger and Johansen tests.

1. Engle-Granger Test:

To examine the presence of cointegration between variables, researchers developed a new method based on testing the null hypothesis that no cointegration exists. This method involves estimating a regression between two variables using ordinary least squares (OLS) and then testing for a unit root in the residuals. If the residuals contain a unit root (i.e., are non- stationary), the null hypothesis of no cointegration is accepted. However, if the residuals are stationary and do not contain a unit root, the null hypothesis is rejected, and the alternative hypothesis of cointegration is accepted. This process is carried out in two stages.

In the first stage, the degree of integration of the variables is tested using unit root tests. For cointegration to exist, the time series must be integrated of the same order. If the time series are integrated of different orders, cointegration is ruled out. However, if the series are non- stationary at their original level but integrated of the first order, the relationship between them can be estimated using OLS after transforming the variables into first differences. If the estimated relationship is significant, it indicates that the relationship is not spurious.

At this point, the stability of the residuals (e_t) is verified using the Dickey-Fuller test, with an equation that excludes a constant and trend, and using MacKinnon critical values. If the null hypothesis is accepted, it is concluded that the residuals are non-stationary, and thus there is no cointegration between the time series variables. If the null hypothesis is rejected, the relationship is cointegrated, indicating long-term equilibrium. This allows for the estimation of an ARDL model to obtain short- and long-term estimates. Subsequently, an Error Correction Model (ECM) is used to estimate short-term parameters, which is the second stage of this process.

Second Stage:

The Error Correction Model (ECM) is estimated by including the lagged residuals in the long- term regression, alongside the differences of the other variables, as per the following formula:

$\Delta Y_t = a_1 \Delta X_t + a_2 e_{t-1} + e_t$......................................(12)

represents the error correction coefficient, which measures the speed of adjustment or feedback through which deviations from equilibrium in the short term are corrected toward the long- term equilibrium level. It is expected to be negative and statistically significant.

Regression Methodology (ARDL):

The tests discussed earlier require that the variables in the study be integrated of the same order, along with a sufficient sample size. Therefore, the Autoregressive Distributed Lag (ARDL) approach to bounds testing has become widely used in recent years due to its advantages, as highlighted by Pesaran and Shin in 1999. The ARDL model has several features, including the lack of a requirement for variables to be integrated of the same order, unbiased parameter estimates, and the ability to address autocorrelation and heteroskedasticity.

The bounds test in ARDL models plays a crucial role in identifying long-term relationships between variables and confirming the presence of cointegration, allowing for the analysis of dynamic relationships. ARDL models are powerful tools for econometric analysis and are widely used in academic research and applications. To conduct a bounds test for cointegration using the ARDL methodology, three steps are required:

Determine the optimal lag length for each variable, reflecting the dynamics of the model, based on the Akaike Information Criterion (AIC).

Test for cointegration using the following equation:

$$\Delta Y_t = a_0 + \sum_{i=1}^{m} \beta_1 \ \Delta Y_{t-i} + \sum_{i=0}^{n} \emptyset_i \Delta X_{t-i} + \lambda_1 Y_{t-i} + \lambda_2 X_{t-1} + \varepsilon_t \ldots\ldots(13)$$

$(2\lambda, 1\lambda)$ represent the coefficients of a long-term relationship, and the bounds test depends on the (F) statistic or the (WALD) statistic through:

- Null hypothesis (H0: $\lambda 1 = \lambda 2 = 0$) i.e. no integration
- Substitution hypothesis ($0 \neq$H1: $\lambda 1 \neq \lambda 2$) i.e. the presence of co-integration

Second: After confirming the validity of the joint integration hypothesis, we estimate a long- run equation, in terms of:

$$Y = \alpha 0 + \sum_{i=1}^{m} a_i Y_{t-i} + \sum_{i=0}^{n} \mathrm{P}_i X + u_t \ldots\ldots\ldots\ldots()$$

(m,n): represent the lag period, ai, $\alpha 0$)): coefficients of the variables and ϑ_i are parameters for estimation while Ui: the random error term

Third Step

The short-term dynamic parameters are estimated by estimating the long-term error correction model, via the following equation:

$$\Delta Y = \chi_0 + \sum_{i=1}^{m} \beta_1 \Delta Y_{t-i} + \sum_{i=0}^{n} \theta_i \Delta X_{t-i} + \lambda ECM_{t-1} + e_t \ldots\ldots\ldots\ldots()$$

Determinants of tourism demand in Tipaza province

To avoid the false regression referred to above in the initial formulation of the model within the framework of short-term analysis, and based on the least squares method, and to estimate the temporal relationship between tourism demand and various independent variables represented in some characteristics of local tourism, many approaches have been adopted to improve the performance of economic measurement models in terms of interpretation and prediction. Since the end of the seventies, dynamic models have appeared in the literature of tourism demand, such as the distributed lag model (MDL), followed by the autoregressive distributed lag model (ARDL) in its traditional form, which was used to absorb the problem of autocorrelation, which was then considered as just noise; to later develop into a very important model in econometric studies, then moved through mathematical transformations to also be expressed by the error correction model (ECM).

The modeling process goes through several stages to reach a dynamic formulation of this type, starting from applying stability and joint integration tests, passing through statistical and econometric tests and reaching the estimation process.

Third: Stability Tests

Since the early 1980s, econometricians have believed in the existence of long-term relationships between variables in applied economic studies. It was assumed that the underlying time series were stationary or at least stable around a deterministic trend, and that they exhibited long-term relationships. It was natural to formulate econometric models using traditional methods, assuming that the means and variances of the variables were constant and independent of time. The estimated models were used to analyse abstractly formulated theories, predict policies, and evaluate and stimulate them. However, recent developments in econometrics have revealed that, in many cases, most time series are non-stationary, contrary to earlier beliefs. Some time series may exhibit a tendency to drift away from their mean over time, while others may converge toward their mean, thus tending toward stability. Classical estimation of variables with such relationships often leads to misleading conclusions or spurious regression.

To overcome the problem of non-stationarity and the a priori constraints on the lag structure in a model, econometric analysis of time series data has increasingly shifted toward the issue of cointegration. This is because cointegration is a powerful tool for detecting the existence of a steady-state equilibrium between variables. Cointegration has become a key requirement for any economic model using non-stationary time series data. If the variables are not cointegrated, we face problems of spurious regression, and the results become almost meaningless.

It is now important to study the stability of time series using modern tests that account for the presence of seasonal components with a periodicity of 12, representing months rather than seasons. It has been shown that the various time series under study contain regular periodic effects. The choice of stability tests is based on this consideration, which is provided by standard software packages like EViews 12 and later versions.

1- Table (3): Stability study

Variable	Components	Seasonality	Critical values 5%	The test	Critical values 5%	Décision
LVT	Constant	18.11	6.89	-2.99	-2.91	I(0)
LNT	Constant, Trend	4.01	1.39	-4.41	-4.13	I(0)
LNLT	Constant, Trend	3.26	1.39	-5.12	-3.50	I(0)
LNCH	Constant	3.32	1.81	-8.03	-2.92	I(1)
LEMP	Constant	21.26	6.85	-8.03	-2.91	I(1)

2- Source: Researcher's numbers from Eviews outputs

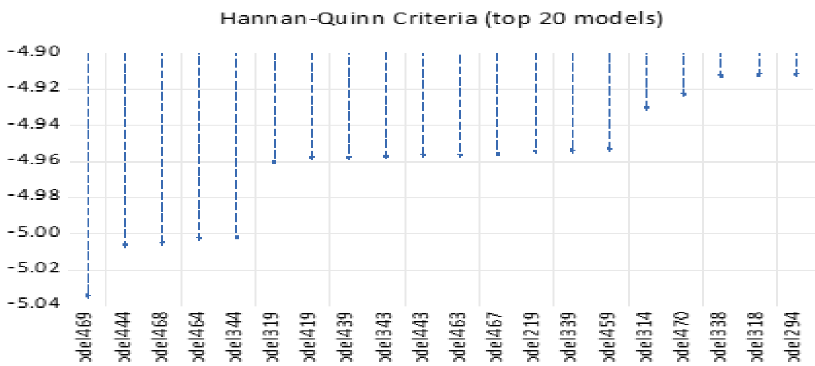
After this description and study of stability, a preliminary ARDL model can be formulated to know the trends of the model under the use of a sample of annual observations and then converting them to monthly via a seasonal coefficient. In terms of significance, the model was acceptable, but it suffers from dynamic parameters' sign. The sign of the (LNCH(-1) and (LEMP(-1) parameters are negative, and not conform to the sense of the dynamic theory, as it is impossible to admit contradiction in the sign of the same dynamic variable estimator, as shown in table (4):

Table (4) ARDL model estimation

Variable	Coefficient	Std. Error	t-Statistic	Prob.*
LNT(-1)	0.4331	0.077	5.6228	0.0000
LNCH	0.7926	0.0326	24.30	0.0000
LNCH(-1)	-0.3606	0.0687	-5.2435	0.0000
LEMP	0.0989	0.021	4.6830	0.0000
LEMP(-1)	-0.08	0.022	-3.9556	0.0000
LNLT	-0.082	0.030	-2.7331	.00000
C	1.7117	0.24	6.97	0.0000
R-squared	0.9943	Mean dependent var		8.9069
Adjusted R-squared	0.99	S.D. dependent var		0.2144
S.E. of regression	0.01	Akaike info criterion		-5.169
Sum squared resid	0.0085	Schwarz criterion		-5.4457
Log likelihood	176.95	Hannan-Quinn criterion		-5.6174
F-statistic	2258	Durbin-Watson stat		1.18
Prob(F-statistic)	0.000000			

Source: Prepared by the researcher based on Eviews 12.

In addition to the above, the following slowdown test table was used:



Source: Eviews extract

In order to determine the validity of the previous estimate, we must research the topics of time series stability and from there move on to studying the extent to which the condition of joint integration is achieved, as follows:

Cointegration Test:

To avoid the spurious regression that often characterizes short-term analysis and to estimate the causal relationships between tourism demand and independent variables, modern econometric techniques have been adopted for modelling tourism demand and predicting its future behavior. Dynamic models have emerged in the literature on tourism demand, such as the Distributed Lag Model (DL), the

Autoregressive Distributed Lag Model (ARDL), and the Error Correction Model (ECM). We adopt the latter formulation after conducting a cointegration test and rejecting the null hypothesis based on the residuals of the estimation, which shows signs of spurious regression.

Test for Stability of Residuals:
Initially, the time series appear generally stable, which is attributed to the modest effects of the general trend and seasonality. We will attempt to confirm or refute this using statistical evidence, which prompts us to consider using one of the stability tests that take into account both the general trend (Dickey-Fuller) and seasonality (HEGY).

Table (5) Stability test

Augmented Dickey-Fuller Unit Root Test on ET

Null Hypothesis: ET has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.326083	0.0205
Test critical values:	1% level	-2.604746	
	5% level	-1.946447	
	10% level	-1.613238	

*MacKinnon (1996) one-sided p-values.

Indeed, after calculating the residual vector (ET) and testing its stability using the Dickey- Fuller test, the time series is found stable, leading to rejection of the null hypothesis. This result is encouraging but requires confirmation through the Bounds Test.

Based on this, and after studying the stability of the variables using the Dickey-Fuller (1979) and HEGY tests to account for seasonal effects in the various time series, the study confirms the stability assumption. The calculated statistics for different seasonal frequencies were greater than the critical values at the 5% level, which we consider quite normal given the limited annual time period. Therefore, the Autoregressive Distributed Lag (ARDL) model can be applied as variables are either I(0) or I(1). This prompts us to conduct the Bounds Test to investigate whether cointegration exists and to select the appropriate modeling approach, whether it be the Dynamic Distributed Lag Model or the Error Correction Model (ECM).

Second Method: Bounds Test
The Bounds Test for cointegration is a valuable tool for assessing whether there is a long- term relationship between two or more time series variables. Cointegration refers to the situation where two or more non-stationary time series variables move together in the long run, even though they may exhibit short-term fluctuations. This implies an equilibrium relationship between the series, indicating that deviations from this equilibrium will eventually be corrected.

Table (6) Bound Test

F-Bounds Test		Null Hypothesis: No levels relationship		
Test Statistic	Value	Signif.	I(0)	I(1)
F-statistic	16.71158	10%	3.17	4.14
k	2	5%	3.79	4.85
		2.5%	4.41	5.52
		1%	5.15	6.36

Source: Researcher's work based on Eviews 12.

If the calculated F-statistic is greater than the critical value, we reject the null hypothesis and accept the presence of cointegration, indicating a long-term relationship. In this case, we accept the existence of a cointegrating relationship because the F-statistic (16.71158) > 4.85, which supports the Dickey-Fuller test and allows for the estimation of an ECM (Error Correction Model). The ECM is a tool for analyzing time series data, as it reflects the long- term shared stochastic trend among the underlying variables.

1. Error Correction Model (ECM):
The term "error correction" refers to the idea that deviations from long-term equilibrium in the previous period influence short-term dynamics, as shown by its coefficient in Table (7), which is negative and highly significant. This coefficient expresses the speed at which the dependent variable returns to equilibrium due to the influence of other independent variables, thereby correcting deviations from equilibrium over time. From this, we conclude that there is a process that prevents economic variables from drifting too far from their long-term equilibrium levels. Thus, this model measures the speed at which short- term imbalances are adjusted and conditions return to long-term equilibrium. Evidence of a long-term equilibrium relationship between the studied variables exists if the error correction coefficient is negative and significant, as shown below.

Table (7) Error Correction Model (ECM) according to the ARDL methodology

ECM Regression				
Case 3: Unrestricted Constant and No Trend				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
D(LNCH)	**0.7928**	**0.018**	**42.99**	**0.0000**
D(LEMP)	**0.09**	**0.019**	**5.17**	**0.0000**
LVT	**0.205**	**0.024**	**8.28**	**0.0000**
CointEq(-1)*	**-0.56**	**0.06**	**-8.30**	**0.0000**

Source: Researcher's work based on Eviews 12.

The last value in the equation represents the series of residuals trailed by one time period, while D here represents the series of differences for each variable. The ECM coefficient estimated above, which is (-0.56), also shows how quickly/slowly the variables return to their equilibrium level in the dynamic model, indicating the possibility of achieving equilibrium in the long run. Thus, the deviation from the long-term equilibrium can be corrected by 56% per month, which we consider fast, as it takes about two months to return to the equilibrium level, given the abundance of local tourism, the state's tourism traditions, and the country's recovery from the Covid-19 shock. The analysis of the short-term estimation results is important, as a 1% increase in the number of rooms contributes to a significant increase in demand for hotels, amounting to 79%, while a 1% increase in employment causes a 9% increase in demand, while an increase in visitors contributes 20% to hotel demand. We note that this model is considered better in terms of explanation than the lag model.

Estimation of long-run relationship

These models are the other side of the same coin, the first side being error correction models. This estimation is preferred in the absence of cointegration, and both have the advantage of being able to overcome the problem of spurious regression and achieve long- run feature estimation, which is very important.

Table (8) Results of the long-term relationship

Levels Equation				
Case 3: Unrestricted Constant and No Trend				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNCH	0.76	0.06	12	0.0000
LEMP	-0.14	0.04	-3.37	0.0014
LNLT	0.02	0.01	1.12	0.2647
C	3.01	0.13	22.46	0.0000
EC = LNT - (-0.76*LNCH -0.14*LEMP -0.02*LNLT +3.02)				

Source: Eviews 12 output.

The transition from the error correction model to the lag model is done using mathematical transformations, and given the availability of the first estimate mentioned above, which produced the short-term parameters in addition to proving the existence of joint integration, in addition to determining the period of return to the equilibrium level within two months. Table (8) shows the long-term relationship between the dependent variable and the independent variables. The parameters of the long-term relationship were conflicting in terms of the sign of the number of workers, which was negative despite its statistical significance, which cannot be easily accepted, especially since the sector is relied upon to increase employment, and this result is not consistent with the short-term result. This result can be interpreted as temporary and seasonal workers having a positive impact on providing tourism services, which are automatically disposed of after the end of the summer season, while permanent workers have a negative impact, which is reflected in the long-term parameters. The number of beds did not have any impact on hotel demand. This may be due to the study sample and the absence of some important variables from the model that may be related to determining tourism demand. In conclusion, we consider that tourism in Tipaza is linked, according to this dynamic study, to the variable "number of rooms", which is linked to the reception capacity, being very limited in the summer and autumn. We also notice through the study the existence of annual local tourism, although limited, particularly as tourism demand is linked to the private sector spread across the entire coast of Tipaza, while the potential of the public sector is concentrated in hotels and some centres in the eastern coastal strip.

Conclusion

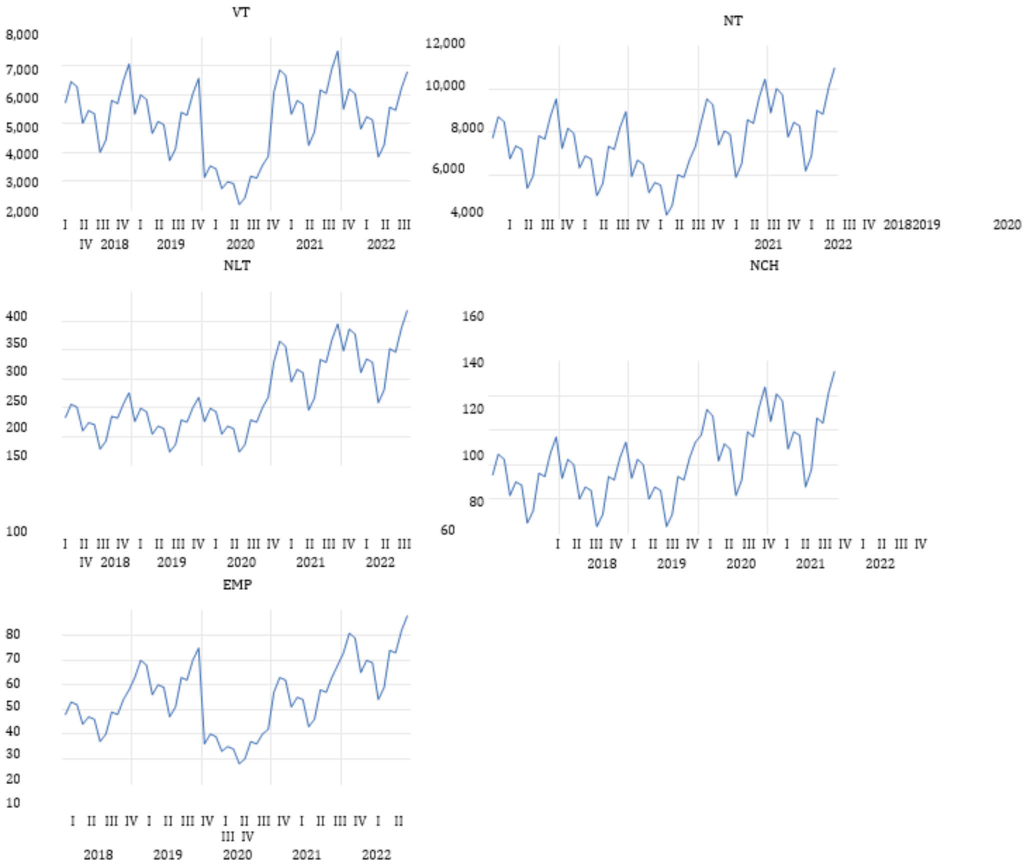
The need to use econometric techniques has become increasingly vital in various fields of scientific research, specifically in studies of tourism demand, with the aim of understanding tourist behaviour and identifying the factors influencing tourism demand. In this context, our study on tourism in Tipaza is conducted under the impact of the COVID-19 pandemic, which partially disrupted domestic tourism activities. These circumstances incited us to investigate the effectiveness of statistical methods in predicting the recovery of tourism demand during sudden and volatile crises, motivating us to adopt appropriate dynamic models to analyse the phenomenon. We seek to identify the most important factors influencing the choice of Tipaza as a local tourist destination. We test a sample of variables, including the number of tourists in the region (VT), the number of hotel nights (NT), the number of beds (NLT), the number of rooms (NCH), and employment (EMP). We also attempt to compare the use of the number of tourists or the number of hotel nights as the dependent variable, with the remaining variables considered as independent.

Econometric analysis relies on a set of requirements and hypotheses related to modelling and estimation, necessitating the study of data from multiple angles. This begins with descriptive statistical analysis to measure the homogeneity of the data and verify whether the time series experience events, shocks, or structural changes that need to be taken into account during the modelling process. This examination includes checking the normality of the variables, as well as graphical analysis to uncover the components of the time series, such as the global trend, seasonality, and randomness. All of these factors may affect the stability of the time series, making their study essential before embarking on any econometric analysis.

The results of the study indicate that the error correction model (ECM) illustrates the relationship between long-term equilibrium and short-term dynamics, as it helps correct deviations and return variables to their equilibrium levels quickly. The estimated time to return to equilibrium was about two months, reflecting the flexibility of tourism activity in Tipaza after the COVID-19 shock. It was also found that the number of rooms plays a major role in hotel demand, as a 1% increase in the number of rooms leads to a 79% increase in demand, while employment contributes less, with a 1% increase in employment raising demand by only 9%. However, the long-term relationship showed a contradiction in the impact of employment, as the effect of permanent employment was negative despite its statistical significance, which may be attributed to the seasonal nature of the sector. Based on these results, we conclude that tourism in Tipaza depends mainly on the number of rooms as a determining factor for demand, especially given the limited reception capacity during peak seasons. Additionally, tourism demand is more influenced by the private sector spread along the coast, while the public sector's capabilities remain concentrated in the eastern coastal strip of the province.

Appendices:

Figure 1: Variables plots



	VT	NT	NLT	NCH	EMP
Mean	5088.550	7549.233	219.8000	102.4000	44.90000
Median	5341.000	7544.000	200.0000	101.0000	44.50000
Maximum	7534.000	10981.00	369.0000	154.0000	78.00000
Minimum	2185.000	4147.000	123.0000	64.00000	18.00000
Std. Dev.	1303.132	1538.857	64.32987	21.08755	14.32752
Skewness	-0.425188	0.004578	0.564176	0.333525	0.156162
Kurtosis	2.278307	2.395956	2.238884	2.609150	2.285334
Jarque-Bera	3.109956	0.912381	4.631194	1.494300	1.520733
Probability	0.211194	0.633693	0.098707	0.473715	0.467495
Sum	305313.0	452954.0	13188.00	6144.000	2694.000
Sum Sq. Dev.	1.00E+08	1.40E+08	244161.6	26236.40	12111.40
Observations	60	60	60	60	60

Source: Eviews output

Dependent Variable: LNT Method: Least Squares Date: 05/18/24 Time: 17:11 Sample: 2018M01 2022M12
Included observations: 60

Variable	CoefficientStd. Error	t-Statistic	Prob.
C	4.7716140.188070	25.37150	0.0000
LNCH	1.0492470.131212	7.996587	0.0000
LNLT	-0.3062840.094249	-3.249733	0.0020
LEMP	0.2507890.032070	7.820106	0.0000

R-squared	0.929762Mean dependent var	8.907686
Adjusted R-squared	0.925999S.D. dependent var	0.212695
S.E. of regression	0.057860Akaike info criterion	-2.797250
Sum squared resid	0.187474Schwarz criterion	-2.657627
Log likelihood	87.91751Hannan-Quinn criter.	-2.742636
F-statistic	247.0958Durbin-Watson stat	0.175045
Prob(F-statistic)	0.000000	

Dependent Variable: LNT Method: Least Squares Date: 05/18/24 Time: 17:29 Sample: 2018M01 2022M12
Included observations: 60

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.984370	0.190520	26.16188	0.0000
LNCH	0.658493	0.056755	11.60240	0.0000
LEMP	0.237094	0.034355	6.901368	0.0000

R-squared	0.916516	Mean dependent var	8.907686
Adjusted R-squared	0.913587	S.D. dependent var	0.212695
S.E. of regression	0.062524	Akaike info criterion	-2.657820
Sum squared resid	0.222829	Schwarz criterion	-2.553103
Log likelihood	82.73460	Hannan-Quinn criter.	-2.616859
F-statistic	312.8826	Durbin-Watson stat	0.214948
Prob(F-statistic)	0.000000		

Quarterly Stability Test Model: Employment Variable

Seasonal Unit Root Test for LEMP Method: Traditional HEGY
Null Hypothesis: Unit root at specified frequency Periodicity (Seasons): 12
Non-Seasonal Deterministic: None Seasonal Deterministic: None
Lag Selection: 0 (Automatic: AIC, maxlags=12) Sample Size: 48

Significance Level

	Test Stat.	1%	5%	10%
Frequency 0 n=40	-0.447233			
		-2.53	-1.88	-1.59
n=60		-2.57	-1.92	-1.60
n=48*		-2.55	-1.90	-1.59
Frequency 2PI/12 and 22PI/12	2.051163			
n=40		30.65	7.98	3.66
n=60		30.93	7.99	3.73
n=48*		30.76	7.98	3.69
Frequency 4PI/12 and 20PI/12 n=40	4.510623			
		30.65	7.98	3.66
n=60		30.93	7.99	3.73
n=48*		30.76	7.98	3.69
Frequency 6PI/12 and 18PI/12 n=40	2.816067			
		30.65	7.98	3.66
n=60		30.93	7.99	3.73
n=48*		30.76	7.98	3.69
Frequency 8PI/12 and 16PI/12 n=40	0.932414			
		30.65	7.98	3.66
n=60		30.93	7.99	3.73
n=48*		30.76	7.98	3.69
Frequency 10PI/12 and 14PI/12 n=40	6.418032			
		30.65	7.98	3.66
n=60		30.93	7.99	3.73
n=48*		30.76	7.98	3.69
Frequency PI n=40	-2.347651			
		-2.53	-1.88	-1.59
n=60		-2.57	-1.92	-1.60
n=48*		-2.55	-1.90	-1.59
All seasonal frequencies n=40	21.56276			
		28.09	7.38	3.43
n=60		28.13	7.36	3.49
n=48*		28.11	7.37	3.45
All frequencies n=40	21.26457			
		25.99	6.89	3.23
n=60		26.14	6.85	3.28
n=48*		26.05	6.87	3.25

*Note: Obtained using linear interpolation.

Dependent Variable: LEMP-LEMP (-12) Method: Least Squares
Date: 05/16/24 Time: 23:13
Sample (adjusted): 2019M01 2022M12

Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
OMEGA(0)	0.000305	0.000683	0.447233	0.6574
OMEGA(2PI/12)	-0.013563	0.028432	-0.477048	0.6362
OMEGA(22PI/12)	-0.053114	0.027173	-1.954663	0.0584
OMEGA(4PI/12)	-0.112546	0.073346	-1.534455	0.1337
OMEGA(20PI/12)	-0.166609	0.069119	-2.410479	0.0212
OMEGA(6PI/12)	-0.132100	0.083056	-1.590490	0.1205
OMEGA(18PI/12)	-0.136354	0.083102	-1.640791	0.1096
OMEGA(8PI/12)	-0.074963	0.063069	-1.188590	0.2424
OMEGA(16PI/12)	-0.041314	0.063863	-0.646916	0.5218
OMEGA(10PI/12)	-0.487825	0.140407	-3.474371	0.0014
OMEGA(14PI/12)	-0.102525	0.156461	-0.655276	0.5165
OMEGA(PI)	-0.249295	0.106189	-2.347651	0.0245

R-squared	0.868521	Mean dependent var	0.123327
Adjusted R-squared	0.828347	S.D. dependent var	0.494873
S.E. of regression	0.205031	Akaike info criterion	-0.118994

ARDL Long Run Form and Bounds Test Dependent Variable: D(LNT)
Selected Model: ARDL(1, 1, 0, 1)
Case 2: Restricted Constant and No Trend Date: 05/18/24 Time: 15:00
Sample: 2018M01 2022M12
Included observations: 59

Conditional Error Correction Regression

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.711737	0.245245	6.979691	0.0000
LNT(-1)*	-0.566807	0.077041	-7.357204	0.0000
LNCH(-1)	0.432068	0.078296	5.518395	0.0000
LNLT**	-0.082592	0.030218	-2.733186	0.0086
LEMP(-1)	0.011695	0.010846	1.078344	0.2860
D(LNCH)	0.792675	0.032608	24.30959	0.0000
D(LEMP)	0.098951	0.021130	4.683036	0.0000
LVT	0.205103	0.028082	7.303571	0.0000

* p-value incompatible with t-Bounds distribution.
** Variable interpreted as Z = Z(-1) + D(Z).

Levels Equation
Case 2: Restricted Constant and No Trend

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNCH	0.762284	0.063496	12.00526	0.0000
LNLT	-0.145715	0.043177	-3.374858	0.0014
LEMP	0.020634	0.018297	1.127697	0.2647
C	3.019967	0.134416	22.46725	0.0000

EC = LNT - (0.7623*LNCH -0.1457*LNLT + 0.0206*LEMP + 3.0200)

F-Bounds Test Null Hypothesis: No levels relationship

Test Statistic	Value	Signif.	I(0)	I(1)
			Asymptotic: n=1000	
F-statistic	12.78849	10%	2.37	3.2
k	3	5%	2.79	3.67
		2.5%	3.15	4.08
		1%	3.65	4.66
			Finite Sample: n=60	
Actual Sample Size	59	10%	2.496	3.346
		5%	2.962	3.91
		1%	4.068	5.25
			Finite Sample: n=55	
		10%	2.508	3.356
		5%	2.982	3.942
		1%	4.118	5.2