*Corresponding Author's Email: danielalika@uamd.edu.al
Proceedings of the International Conference on Research in Social Sciences
Vol. 2, Issue. 1, 2025, pp. 1-12
DOI: https://doi.org/10.33422/rssconf.v2i1.868

ISSN: 3030-0452 online

A Structural Equation that Estimates the Empact of GDP and Some of Its Factors on Tourism

Dr. Daniela Lika, and Dr. Jonida Gashi
Department of Economics, University Aleksandër Moisiu, Durres, Albania

Abstract

This paper investigates the relationship between economic activity and some of its factors such as trade and capital, and Tourism. The interest of this study represents a "system of relationships", some independent variables and some dependent variables as well When researchers aim to model a complex causal system with queries that qualify as a "system" of relationships, structural equation modeling (SEM) proves to be very useful. For this reason, structural equation modeling is employed in this study. The analysis examines European countries. The data of the variables used are data belonging to European countries. The structural equation model estimated that the improvement of infrastructure and innovation have a significant effect respectively on trade and capital, which are two crucial factors that affect the economic growth of countries. The model also estimates that economic growth significantly impacts the growth of tourism. Therefore, it is essential to emphasize that the use of initiatives to improve infrastructure, innovation, and economic growth would boost the development of Tourism in countries.

Keywords: GDP, Gross Capital Formation, SEM, Tourism, Trade

1. Introduction

The outbreak of the COVID-19 pandemic seems to brought global tourism to a standstill in the years following the pandemic. The United Nations World Tourism Organization's report brings good news for global tourism. Around 1.1 billion tourists traveled internationally in the first nine months of 2024, as the global tourism sector recovered 98% of pre-pandemic levels. According to the latest World Tourism Barometer by UN Tourism, a full recovery from the biggest crisis in the sector's history is expected by the end of the year (UNWTO, 2024).

The Covid-19 pandemic also affected the global GDP, which suffered a contraction in 2020, with GDP falling to \$82.68 trillion from \$85.13 trillion in 2019. However, global GDP recovered in 2021 to \$87.93 trillion, \$90.77 and \$93.35 trillion in 2022 and 2023 respectively.

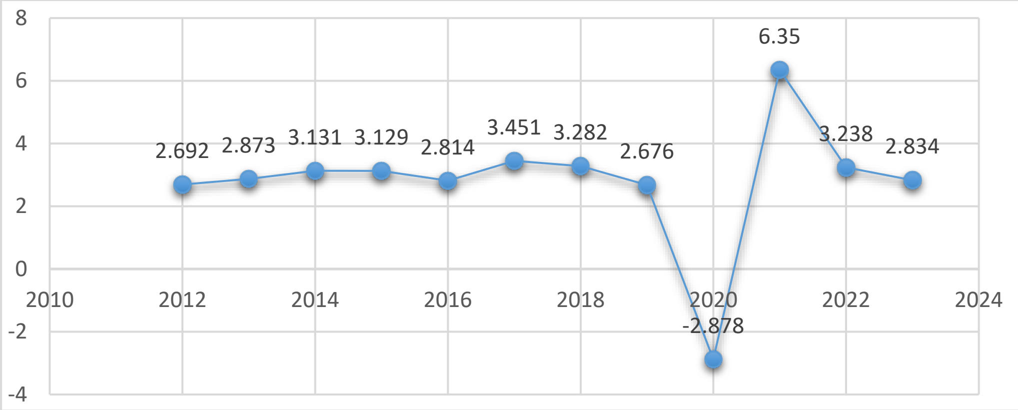

Figure 1: World GDP growth (annual %)
Source: Author's processing, data from World Bank

There have been many studies that have analyzed the relationship between economic growth and tourism. In 2015, a bidirectional causality relationship between tourism and economic growth was the main finding in a study of 12 Mediterranean nations. In order to achieve high economic growth, policy-makers should focus on developing the tourism sector (Bilen et al., 2015). In a paper focusing on the case of Romania, Badulescu et al. (2020) demonstrated a bidirectional causal relationship between GDP and the number of international tourist arrivals. Many other studies have examined the impact that tourism has on economic growth and vice versa. Naseem (2021) investigated the role of tourism in promoting economic growth and found a strong link. The role of tourism in the economic development has also been examined in a paper that includes long-term data of the Spanish economy. The results indicate that, at least during the last three decades, economic growth in Spain has been sensitive to persistent expansion of international tourism (Balaguer & Cantavella, 2022). According to literature reviews economic growth and tourism are interconnected. Therefore, an increase in the economic activity will boost tourism sector too. Measuring the effect that a country's GDP can have on its tourism would be of great value for policy suggestions, given that tourism is one of the fastest expanding sectors of the world economy in recent years. Since a country's GDP is affected by several variables, this paper is an attempt to reflect this influence too. As a result, we investigate in this paper how GDP and some of its factors are linked to tourism.

2. Methodology

The main objective of the study is to assess the impact of GDP on Tourism. This study also aims to evaluate the impact of factors affecting a country's economic activity (GDP), such as Trade and Gross Capital Formation. Another important goal of this paper is to assess the impact of Logistics Performance of a country on the country’s Trade and the assessment of Innovation of a country on the country’s Gross Capital Formation.

Thus, the purpose of this paper is to explore and elucidate the complex relationships between economic activity, expressed through Gross Domestic Product (GDP), and its influencing factors - specifically Trade, Gross Capital Formation (GCF), Logistics Performance, and Innovation - and how these elements collectively impact the tourism sector in European countries.

As stated above, the interest of this study represents a "system of relationships", some independent variables and some dependent variables as well. When researchers aim to model a complex causal system with queries that qualify as a "system" of relationships, structural equation modeling (SEM) proves to be very useful. For this reason, SEM is used in this study. Initially, the diagram of the structural equation is modeled according to the purposes of the study and the model is further evaluated and analyzed, in an attempt to explain that factors such as Logistics Performance and Innovation affect a country's Trade and Gross Capital Formation, respectively, and the latter's affect a country's GDP, thus ultimately affecting a country's Tourism. Thus, hypotheses of the study can be summarized as follows:

- Hypothesis 1: Logistics Performance positively affects Trade.
- Hypothesis 2: Innovation positively affects Gross Capital Formation (GCF).
- Hypothesis 3: Trade and Gross Capital Formation positively affect economic growth (GDP).
- Hypothesis 4: Economic growth (GDP) positively affects Tourism.

Data from forty-six European countries for 2022 were used for each of the six variables, organized in a cross-sectional. The countries under consideration are: Albania, Armenia, Austria, Azerbaijan, Belgium, Bulgaria, Bosnia and Herzegovina, Belarus, Switzerland, Cyprus, Czechia, Germany, Denmark, Spain, Estonia, Finland, France, United Kingdom, Georgia, Greece, Croatia, Hungary, Ireland, Iceland, Italy, Kazakhstan, Lithuania, Luxembourg, Latvia, Moldova, North Macedonia, Malta, Montenegro, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Rwanda, Serbia, Slovak Republic, Slovenia, Sweden, Turkey, Ukraine. Data for countries with missing information were excluded. For the purpose of comparative analysis over the years to observe the sustainability of the model and its relationships, it is remodeled and evaluated using data from 2018 and 2012 from the same 46 European countries. The analysis and evaluation of the model generate results that serve as the basis for recommendations and guiding conclusions regarding economic activity (GDP) and Tourism.

Table 1: Sources of data used in the model

Variable	Definition	Measure	Source
Tourism	Total Tourism Expenditure	(Current US$)	UNWTO https://www.unwto.org/
GDP	Economic Activity	(Current US$)	The World Bank https://data.worldbank.org/
Trade	Exports and Imports	(Current US$)	The World Bank https://data.worldbank.org/
GCF	Gross Capital Formation	(Current US$)	The World Bank https://data.worldbank.org/
Logistics	Logistics Performance Index	Index	The World Bank https://lpi.worldbank.org/
Innovation	Global Innovation Index	Index	WIPO https://www.wipo.int/en/web/global-innovation-index

Source: Author's illustrations

2.1 Model Construction

After reviewing the supporting literature and describing all variables designed for use in the model, as well as the relationships between them, the general functional form of the model is presented, and the model is graphically constructed. The equation of the model is:

$$Tourism_i{}^1 = f\{GDP_i{}^2[GCF_i{}^3(GII_i{}^4)Trade_i{}^5(LPI_i{}^6)]\} \quad (1)$$

The model diagram is presented in Figure 2. This model will be analyzed, and after evaluating the relationships between variables, their significance, and the overall fit of the model, a decision will be made regarding its appropriateness or potential modifications. The structuring of the model (including the construction of the diagram and measurements) is carried out using AMOS software. In Figure 2, the variables are represented within boxes, totaling six variables. Single-headed arrows indicate causal relationships, pointing from independent variables to dependent variables. Double-headed arrows indicate covariances between variables. There are five one-way relationships that illustrate the effects of the independent variables on the dependent variables, as well as four error terms associated with the dependent variables. The model employs the variables described above, using cross-sectional data from 46 European countries for the year 2022. For comparative analysis, data from the same countries are also collected for the years 2018 and 2012.

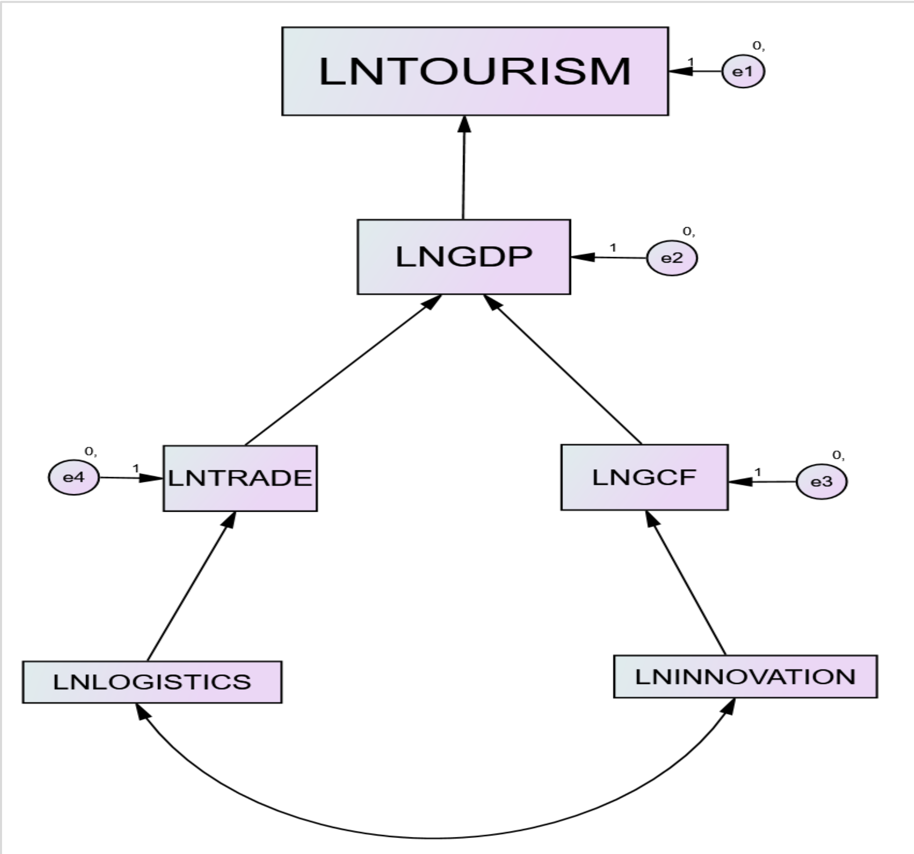

Figure 2: Diagram of structural equation modeling
Source: Author's processing generated through AMOS software

[1] Tourism - Dependent variable of independent variable GDP
[2] GDP - Dependent variable of two independent variables GCF and Trade
[3] GCF - Dependent variable of independent variable GII
[4] GII - Innovation
[5] Trade - Dependent variable of independent variable LPI
[6] LPI - Logistics

2.2 Evaluation Methods

Normality tests: Structural equation modeling requires data normalization. Data normalization was done with SPSS software. To test the normalization of the data were examined: Kolmogorov-Smirnov and Shapiro-Wilk tests, Skewness and Kurtosis (z-values) and Histograms.

Model Fit: Once the model structure is built, it is essential to assess the Model Fit. It refers to how well a proposed model represents the data collected. This section covers the most widely reported fit indices, categorized into three types:

- Absolute model fit

Chi-square (χ^2): It is a single number that indicates whether there is a significant difference between the observation and the expectation. Chi-square values should not be too high compared to the number of degrees of freedom. The chi-square p-value should be greater than 0.05.

RMSEA - Root Mean Square Error of Approximation: For absolute fit, the value of this indicator should be less than 0.08, but even a value not too far from 0.08 (for example 0.1) does not indicate that the model should be excluded (Kumar, 2015).

GFI - Goodness of Fit Index: For values greater than 0.9 the model fit is acceptable (Byrne, 1994).

- Incremental model fit

AGFI - Adjusted Goodness of Fit: For values greater than 0.9 the model fit is acceptable (Byrne, 1994).

CFI - Comparative Fit Index: Values closer to 1 indicate better fit, however values greater than 0.9 are considered a model with acceptable fit (Byrne, 1994).

NFI - Normed Fit Index: For values greater than 0.9 the model fit is acceptable (Byrne, 1994).

TLI Tucker Lewis Index: For values greater than 0.9 the model fit is acceptable (Byrne, 1994).

- Parsimonious model fit

Chi-Square/Degree of freedom (χ^2/df): This ratio should be less than 5 for acceptable model fit (Schumacker & Lomax, 2004).

In Structural Equation Modeling (SEM), there is no strict rule for how many fit indices need to be met for a model to be considered a good fit. However, a common practice is to evaluate multiple fit indices together to get a comprehensive understanding of the model's performance. When the model does not fit well but the values from the fit indices are close to those required for acceptance, modification indices can be used. The model may need to be modified to improve its fit, and the modification indices indicate potential changes that can be made. In AMOS, modification indices are concerned with adding additional covariances within a construct's indicators or relationship paths between constructs (Collier, 2020).

3. Analysis of Results

3.1 Key Model Results for 2022 Data

Normality tests

The data for all variables used in the paper for each year exhibit a normal distribution, as their respective p-values are greater than 0.05. Table 2 summarizes the results of the normality test.

Table 2: Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk			Distribution
	Statistic	*df*	*Sig.*	*Statistic*	*df*	*Sig.*	
Tourism	.088	46	.200*	.962	46	.143	Normal
GDP	.087	46	.200*	.969	46	.262	Normal
Trade	.074	46	.200*	.973	46	.360	Normal
GCF	.116	46	.143	.959	46	.107	Normal
Innovation	.084	46	.200*	.973	46	.351	Normal
Logistics	.115	46	.159	.937	46	.015	Normal

Source: Author's processing generated through SPSS software

Model Fit

To assess the model fit, eight indices are analyzed. The first column of Table 3 presents the eight indicators evaluated for model fit. The second column shows the acceptable values for each indicator for the model to be considered fit. The third and fourth columns present the results: the resulting value of each indicator and the conclusion regarding whether each indicator indicates fitness or unfitness.

Table 3: Model Fit Indices

Indices	**Acceptable value of fit**	**Result value**	**Conclusion**
Chi-Square χ^2, DF, p-value	Not too great and not too far from DF, p-value > 0.05	121.847, 9, 0.000	Unfit
RMSEA	< 0.08	.528	Unfit
GFI	>0.9	.583	Unfit
AGFI	>0.9	.764	Unfit
CFI	>0.9	.761	Unfit
NFI	>0.9	.750	Unfit
TLI	>0.9	.602	Unfit
χ^2/DF	<5	13.539	Unfit

Source: Author's processing generated through AMOS software

The results indicate that the model is not fit, however, the resulting values of the indicators are close to the threshold for acceptance. We can consult the modification indices, which indicate the parameters that should be added to the model to improve its fit. Table 4 presents the data suggested by the modification indices, illustrating the correlations that can be assessed between variables or between variables and error terms.

Table 4: Modification Indices

			M.I.	Par Change
e4	<-->	e3	31.983	1.296
e3	<-->	Lnlogistics	4.294	.071

Source: AMOS software

The suggested correlations include the addition of the covariances e4-e3 and e3-LnLogistics. Once these covariances are added to the model (as shown in Figure 3), the model fit is retested.

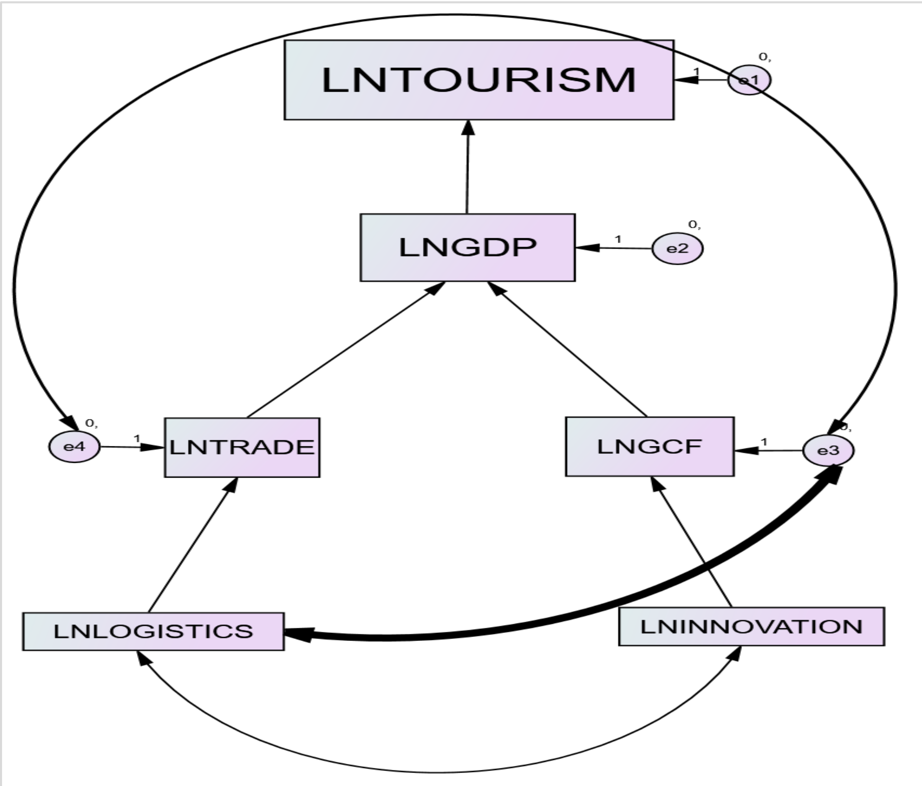

Figure 3: Diagram of structural equation modeling, covariances e4-e3 and e3-LnLogistics are added
Source: Author's processing generated through AMOS software

The results are presented in Table 5. Notably, out of the eight indicators, six indicate fit. This means there are no significant differences between the observed and expected values, allowing us to proceed with the other estimates.

Table 5: Model Fit Indices, after covariances e4-e3 and e3-LnLogistics are added

Indices	**Acceptable value of fit**	**Result value**	**Result**
Chi-Square χ^2, DF, p-value	Not too great and not too far from DF, p-value > 0.05	18.495, 7, .010	Unfit
RMSEA	< 0.08	.191	Unfit
GFI	>0.9	.919	Fit
AGFI	>0.9	.976	Fit
CFI	>0.9	.976	Fit
NFI	>0.9	.962	Fit
TLI	>0.9	.948	Fit
χ^2/DF	<5	.642	Fit

Source: Author's processing generated through AMOS software

Regression Weights

After evaluation, we arrived at a fit model. In this section, the regression weights will be estimated. This analysis helps determine whether the independent variables explain the dependent variable and their significance in explaining the dependent variable. Table 6 presents the results. All independent variables are found to be explanatory and important for their respective dependent variables, as their p-values are less than 0.05.

Table 6: Regression Weights Estimates

			Estimate	S.E.	C.R.	P	Label
LNTRADE	<---	LNLOGISTICS	7.029	1.094	6.425	***	par_1
LNGCF	<---	LNINNOVATION	3.015	.669	4.507	***	par_2
LNGDP	<---	LNGCF	.844	.066	12.887	***	par_7
LNGDP	<---	LNTRADE	.167	.069	2.411	.016	par_8
LNTOURISM	<---	LNGDP	.670	.070	9.574	***	par_5

Source: Author's processing generated through AMOS software

Squared Multiple Correlations

The squared multiple correlation is a measure of how well a given variable can be predicted by a set of other variables. The multiple correlation coefficient is used to assess the quality of prediction of the dependent variable based on the independent variables. It can also be interpreted as the percentage of variance in the dependent variable explained by the independent variables (Abdi, 2007). A high value of the multiple correlation coefficient indicates better predictability. This paper's model includes four dependent variables: Tourism, GDP, Trade, and GCF, for which the multiple correlation coefficients will be estimated. From the results of this test, we will assess whether:

- Tourism is well explained by GDP.
- GDP is well explained by Trade and GCF.
- GCF is well explained by Innovation.
- Trade is well explained by Logistics.

Table 7: The estimates of Squared Multiple Correlations

	Estimate
LNGCF	.277
LNTRADE	.478
LNGDP	.987
LNTOURISM	.671

Source: Author's processing generated through AMOS software

Table 7 shows the results of the squared multiple correlations. The multiple correlation coefficient for Tourism is 0.671, indicating that GDP explains Tourism well, specifically, 67.1 percent of the variance in the Tourism variable is accounted for by GDP.

The multiple correlation coefficient for GDP is 0.987, suggesting that GCF and Trade explain GDP quite effectively, with 98.7 percent of the variance in the GDP variable accounted for by these two factors.

For Trade, the multiple correlation coefficient is 0.478, meaning that 47.8 percent of the variance in Trade is explained by Logistics in this model.

The multiple correlation coefficient for GCF is equal to 0.277. Innovation explains GCF, as 27.7 percent of the variance in the GCF variable in this model is explained by Innovation.

All four dependent variables used in the model are well explained by the corresponding set of independent variables.

3.2 Results for Modeling with 2018 and 2012 Data in the Context of a Comparative Analysis

Model Fit

To conduct a comparative analysis over time and assess the stability of the model, this section analyzes a model constructed similarly to the previous one, using the same variables and relationships but with data from 2018. The model is then re-modeled using data from 2012. After constructing the model shown in Figure 4, which utilizes normalized data for each variable, the model fit is tested. The diagrams of the structural equation models are provided after consulting the modification indices.

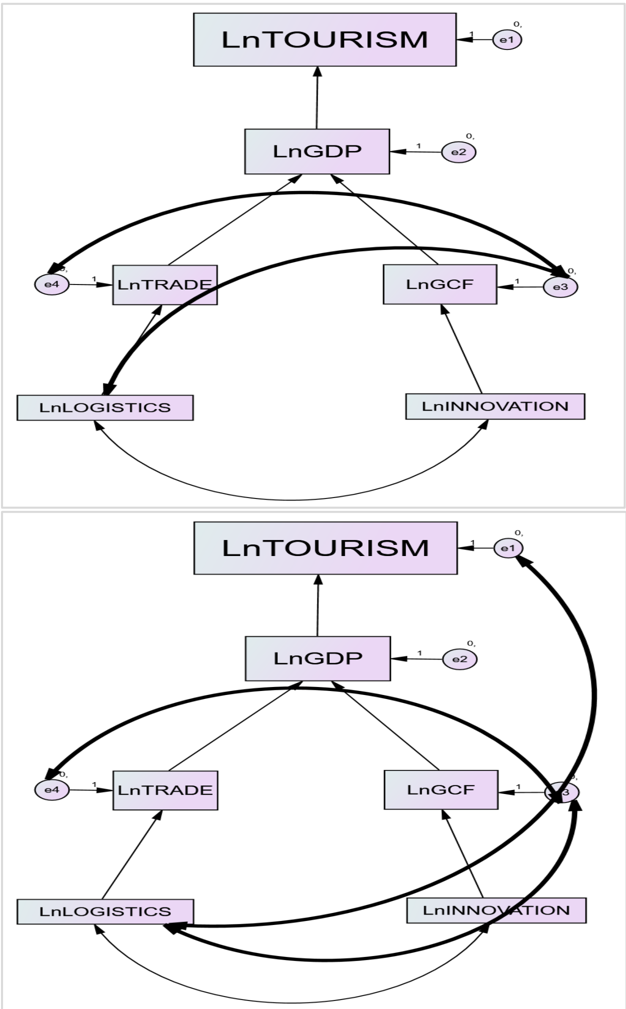

Figure 4: Diagram of structural equation modeling, year 2018 and 2012 respectively
Source: Author's processing generated through AMOS software

The estimates shown in Table 8 for both the 2018 and 2012 data were obtained after consulting the modification indices.

For the model using 2018 data, it is noted that out of eight indicators, seven of them indicate fit. This means that there are no significant differences between the observed and expected values. Therefore, the model is considered fit, allowing us to proceed with the other estimates.

For the model using 2012 data, six out of eight indicators also express fit, and this model is similarly considered fit.

Table 8: Model Fit Indices, 2018 and 2012 Data

Indices	Acceptable value of fit	Result value 2018 Data	Result 2018 Data	Result value 2012 Data	Result 2012 Data
Chi-Square χ^2, DF, p-value	Not too great and not too far from DF, p-value > 0.05	10.582, 7, 0.158	Fit	15.348, 6, .018	Unfit
RMSEA	< 0.08	.107	Unfit	.186	Unfit
GFI	>0.9	.957	Fit	.925	Fit
AGFI	>0.9	.993	Fit	.982	Fit
CFI	>0.9	.993	Fit	.981	Fit
NFI	>0.9	.980	Fit	.970	Fit
TLI	>0.9	.985	Fit	.953	Fit
χ^2/DF	<5	1.512	Fit	2.558	Fit

Source: Author's processing generated through AMOS software

Regression Weights

The following analysis involves the estimation of the regression weights. All independent variables are explanatory and significant for their corresponding dependent variables, as all p-values are less than 0.05. This is true for both models, as Table 9 show.

Table 9: Regression Weights Estimates, 2018 and 2012 Data

			2018 Data					2012 Data				
			Estimate	S.E.	C.R.	P	Label	Estimate	S.E.	C.R.	P	Label
LNTRADE	<---	LNLOGISTICS	8.184	0.945	8.657	***	par_1	7.725	1.03	7.438	***	par_1
LNGCF	<---	LNINNOVATION	4.363	0.850	5.134	***	par_2	3.522	.929	3.792	***	par_2
LNGDP	<---	LNGCF	0.862	0.053	16.418	***	par_5	0.826	.056	14.644	***	par_4
LNGDP	<---	LNTRADE	0.156	0.054	2.878	.004	par_6	0.192	.058	3.304	***	par_5
LNTOURISM	<---	LNGDP	0.707	0.052	13.580	***	par_4	0.738	.050	14.605	***	par_3

Source: Author's processing generated through AMOS software

Squared Multiple Correlations

Table 10 presents the results of the squared multiple correlations for the 2018 and 2012 data. The estimates based on the multiple correlation coefficients for both models indicate that:

- Tourism is well explained by GDP.
- GDP is well explained by Trade and GCF.
- GCF is well explained by Innovation.
- Trade is well explained by Logistics.

Table 10: The estimates of Squared Multiple Correlations, 2018 and 2012 Data

	Esti mate 2018	Estimate 2012
LNGCF	.330	.189

	Estimate 2018	Estimate 2012
LNTRADE	.625	.548
LNGDP	.992	.989
LNTOURISM	.804	.827

Source: Author's processing generated through AMOS software

4. Conclusion

The study conducted using structural equation modeling (SEM) highlights the interconnected between GDP, its contributing factors, and the tourism sector across European countries.

The findings indicate that logistics improvements and innovation significantly influence trade and GCF, both of which are critical components of economic growth (GDP). This highlights the importance of increasing logistic performance and fostering innovation to stimulate economic activity.

The analysis reveals that trade and GCF together account for 98.7% of the variance in GDP. This suggests that policies aimed at improving trade and capital investment can significantly bolster GDP, which boosts tourism.

Economic activity (GDP) has a positive effect on tourism development. The model estimates that GDP explains approximately 67.1% of the variance in tourism expenditure, indicating that as economic activity increases, tourism is likely to expand as well.

The comparative analysis using data from 2018 and 2012 demonstrates the model’s stability over time, suggesting that the relationships among the variables remain stable.

The results suggest policy-makers to prioritize initiatives focused on improving logistics, encouraging innovation, and enhancing trade as a means to stimulate economic growth, which boost the tourism industry.

References

Abdi, H. (2007). *Multiple Correlation Coefficient.* In N. J. Salkind (Ed.), Encyclopedia of measurement and statistics (pp. 593–597). Sage.

Badulescu, A., Badulescu, D., Simut, R., & Dzitac, S. (2020). Tourism—Economic growth nexus. The case of Romania. *Technological and Economic Development of Economy*, 26(4), pp. 867–84. https://doi.org/10.3846/tede.2020.12532

Balaguer, J., & Cantavella, M. (2002). Tourism as a Long-Run Economic Growth Factor: The Spanish Case. *Applied Economics*, 34 (7). https://doi.org/10.1080/00036840110058923

Bilen, M., Yilanci, V., Eryuzlu, H. (2015). Tourism development and economic growth: a panel granger causality analysis in the frequency domain. *Current Issues in Tourism*, 20(1), pp. 27–32. https://doi.org/10.1080/13683500.2015.1073231

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS Windows: Basic concepts, applications, and programming*, (1st ed.). Sage.

Collier, J. (2020). Applied structural equation modeling using AMOS (1st ed.). Routledge. https://doi.org/10.4324/9781003018414-1

Cornell University, INSEAD, & WIPO. (2019). The Global Innovation Index 2019: Creating Healthy Lives—The Future of Medical Innovation, Ithaca, Fontainebleau, and Geneva. Available: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2019.pdf

Cornell University, INSEAD, & WIPO. (2013). The Global Innovation Index 2013: The Local Dynamics of Innovation, Geneva, Ithaca, and Fontainebleau. Available: https://ww.wipo.int/edocs/pubdocs/en/economics/gii/gii_2013.pdf

Deviant, S. (2010). *The Practically Cheating Statistics Handbook.* CreateSpace Independent Publishing Platform

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: *Guidelines for Determining Model Fit*, Electronic Journal of Business Research Methods, 6(1), 53–60.

Keith, T. Z. (2014). *Multiple Regression and Beyond: An Introduction to Multiple Regression and Structural Equation Modeling (2nd ed.).* Routledge. https://doi.org/10.4324/9781315749099

Li, P., Xu, C., & Zhou, Q. (2024). Research on Environmental Responsibility Behavior of Tourists in Seaside Ecological Scenic Spots: The Case of Shenzhen Bay Park. Applied Mathematics and Nonlinear Sciences, 9 (1). DOI: https://doi.org/10.2478/amns-2024-2089

Naseem, S. (2021). The Role of Tourism in Economic Growth: Empirical Evidence from Saudi Arabia. *Economies*, 9 (3). https://doi.org/10.3390/economies9030117

Sturgis, P. (2016). Structural Equation Modelling. National Centre for Research Methods.

UNWTO. (2024). Global tourism set for full recovery by end of the year with spending growing faster than arrivals. Available: https://www.unwto.org/news/global-tourism-set-for-full-recovery-by-end-of-the-year-with-spending-growing-faster-than-arrivals

Wijesekara, C., Tittagalla, C., Jayathilaka, A., Ilukpotha, U., Jayathilaka, R., Jayasinghe, P. (2022). Tourism and economic growth: A global study on Granger causality and wavelet coherence., *PLOS ONE,* 17(9), e0274386. https://doi.org/10.1371/journal.pone.0274386

World Bank. (n.d.). Open data. https://data.worldbank.org/

World Intellectual Property Organization (WIPO). (2023). Global Innovation Index 2023: Innovation in the face of uncertainty. Geneva: WIPO. DOI:10.34667/tind.48220. Available: https://www.wipo.int/documents/d/global-innovation-index/docs-en-wipo-pub-2000-2023-en-main-report-global-innovation-index-2023-16th-edition.pdf

Appendix

Data 2022

Data 2018

Data 2012