

Prof. dr. sc. Vlasta Bahovec
Prof. dr. sc. Ksenija Dumičić
Anita Čeh Časni, dipl. oec.

**MODELIRANJE TURISTIČKE POTRAŽNJE REPUBLIKE
HRVATSKE MODELOM VIŠESTRUKNE LINEARNE REGRESIJE**

**MODELLING CROATIAN TOURISM DEMAND USING
MULTIPLE LINEAR REGRESSION METHOD**

SAŽETAK: Cilj rada je prikazati ekonometrijski model hrvatske turističke potražnje za razdoblje od prvog tromjesečja 1998. do četvrtog tromjesečja 2007. godine, te istražiti koji faktori imaju najveći utjecaj na turističku potražnju aproksimiranu ukupnim brojem ostvarenih noćenja. Turistička potražnja modelirana je na uobičajen način, metodom najmanjih kvadrata (OLS). Rezultati analize sugeriraju da najveći utjecaj na turističku potražnju imaju realni dohodak turista i sezonski faktori.

KLJUČNE RIJEČI: turistička potražnja, model višestruke linearne regresije, metoda najmanjih kvadrata (OLS), ADF test, multikolinearnost.

ABSTRACT: The purpose of this paper is to present an econometric model of Croatian tourism demand for the period between first quarter 1998 and fourth quarter 2007, and to investigate which factors have the most influence on tourism demand approximated by total tourist nights. In order to achieve that, Multiple Regression analysis was applied. Due to heteroskedasticity, tourism demand was modeled using GARCH (1;0). The results of the analysis suggest that the most important determinants of tourism demand in Croatia are real tourist revenues and seasonal factors.

KEY WORDS: tourism demand, multiple linear regression, Augmented Dickey Fuller test (ADF), multicollinearity, heteroscedasticity

1. UVOD

U Republici Hrvatskoj značajan dio prihoda dolazi od turizma, pa je za nositelje ekonomske politike izuzetno bitno poznavati čimbenike koji utječu na turističku potražnju. Budući da je ukupan broj noćenja turista (rezidenata i nerezidenata) glavni indikator razvoja turizma¹, cilj ovoga rada je testirati adekvatnost jednog ekonometrijskog modela hrvatske turističke potražnje za razdoblje od prvog tromjesečja 1998. do četvrtog tromjesečja 2007. godine, te istražiti koji čimbenici imaju najveći utjecaj na turističku potražnju aproksimiranu ukupnim brojem ostvarenih noćenja.

Hrvatska je kao turističko odredište ili destinacija tradicionalno zanimljiva turistima s područja zapadne Europe, naročito Nijemcima, ali i susjednim Slovencima. U prilog tome govore i podatci predloženi u tablici 1.

Tablica 1. Glavne emitivne zemlje prema ostvarenim noćenjima turista u 2007. godini

	Noćenja (u tis.)	Udio u ukupnim noćenjima u %	Stopa prom- jene 2007/2006.	Prosječno trajanje bo- ravka
Njemačka	10.849	19,4	-1,3	7,0
Slovenija	5.690	10,2	8,5	5,6
Italija	5.452	9,7	-0,4	4,4
Češka	4.395	7,8	12,1	6,6
Austrija	4.245	7,6	4,3	5,1
Nizozemska	2.030	3,6	4,8	7,7
Mađarska	1.985	3,5	-9,6	5,2
Poljska	1.834	3,3	3,8	5,7
Slovačka	1.832	3,3	28,2	6,5
Francuska	1.665	3,0	-2,5	3,5
Ostale strane zemlje	9.599	17,1	13,7	4,3
Ukupno strani turisti	49.575	88,5	5,4	5,3
Domaći turisti	6.431	11,5	7,4	3,5
Ukupno	56.005	100,0	5,7	5,0

Izvor: Hrvatski turizam u brojkama, broj 4, 2007

Deset najznačajnijih emitivnih zemalja dalo je 71% ukupno ostvarenih noćenja u 2007. godini, pri čemu su Nijemci ostvarili 19,4%, a Slovenci 10,2% u ukupnom broju noćenja. Najviše stope rasta noćenja u usporedbi s 2006. godinom zabilježene su na slovačkom (28,2%), češkom (12,1%) i slovenskom tržištu (8,5%), dok su tržišta Mađarske (-9,6%) i Francuske (-2,5%) bilježila najveći pad. Najveće prosječno trajanje boravka, približno 7 noćenja, ostvarili su turisti iz Njemačke i Nizozemske, a u smještajnim su objektima, osim domaćih turista, najkraće boravili Francuzi.

¹ Prema: Eurostatu, Panorama on tourism, 2008 edition.

Ako se analiziraju zemlje članice Europske unije u terminima ukupno ostvarene međunarodne turističke potrošnje, prema podacima Eurostata za 2006. godinu, na prvome je mjestu Njemačka s obujmom potrošnje gotovo 59,6 bilijuna eura, slijedi Velika Britanija s 50,3 bilijuna eura, te Francuska čiji je obujam međunarodne turističke potrošnje 2006. godine iznosio 24,8 bilijuna eura. Na četvrtom je mjestu Italija, koju slijede Belgija i Nizozemska. Navedene zemlje članice pripadaju i skupini glavnih emitivnih zemalja prema ostvarenim noćenjima (tablica 1).

Općenito se literatura koja obrađuje turističku potražnju, usredotočuje na analiziranje učinaka raznih odrednica i/ili na predviđanje buduće turističke potražnje. Ovaj rad pripada skupini radova koji se prvenstveno bave determinantama u podlozi turističke potražnje.² U većini postojećih empirijskih studija iz ovog područja korišteni su, kao zavisne varijable, dolasci/odlasci turista i prihod od turizma. Broj ostvarenih noćenja i prosječno trajanje boravka, također su korišteni, ali u mnogo manjoj mjeri. Pri odabiru nezavisnih varijabli modeli turističke potražnje oslanjaju se na teoriju potrošača, koja predviđa da stupanj potrošnje ovisi o potrošačevu dohotku, cijeni dobra ili usluge od interesa, te cijeni njenih supstituta ili komplementa. Kao rezultat toga, dohodak i cijene su najčešće korištene varijable jer najznačajnije utječu na turističku potražnju.

Budući da se potrošnja za turističke usluge povezuje s luksusnim dobrima³, za koja se troši dohodak koji preostane pošto se zadovolje osnovne potrebe, mnogi se istraživači koriste nominalnim ili realnim (*per capita*) dohotkom te bruto domaćim proizvodom kao mjerom dohotka emitivnih zemalja.

Osim što su turisti osjetljivi na vlastiti dohodak, osjetljivi su i na cijene u turističkim destinacijama. U literaturi se za aproksimaciju troška boravka u turističkoj destinaciji najčešće koristi indeks potrošačkih cijena (*Consumer Price Indeks* ili *CPI*), a kao negativna strana takve aproksimacije može se navesti da se troškovi turista i stanovnika destinacija mogu u mnogočemu razlikovati, s obzirom na posebnosti potrošnje turista.

Konačno, model turističke potražnje⁴ općenito se procjenjuje kao funkcija dohotka emitivne zemlje, relativnih cijena, kamatnih stopa, troškova transporta između emitivne zemlje i turističke destinacije kao i kvalitativnim (*dummy*) varijablama i determinističkim trendom kojima se opisuje utjecaj sezone i osnovna razvojna tendencija turističke potražnje.

Ovim se radom modeliranje turističke potražnje u Republici Hrvatskoj nastoji provesti aproksimacijom turističke potražnje ukupnim brojem ostvarenih noćenja za razdoblje od prvog tromjesečja 1998. do četvrtog tromjesečja 2007. godine, pri čemu je korišten model višestruke linearne regresije.

Rad se sastoji od četiri dijela. Nakon uvoda, u drugome su dijelu rada opisani podatci i metodologija istraživanja, u trećem su dijelu iznesene teorijske pretpostavke korištene metode i rezultati analize, a četvrti dio rada donosi zaključke i preporuke za daljnja istraživanja.

² Mervar, A., Payne, J.E.: An Analysis of Foreign tourism Demand for Croatian Destinations: Long-Run Elasticity Estimates, EIZ-WP-0701.

³ Ibid.

⁴ Usporediti s: Witt, S. F. and Witt, C. A., Forecasting tourism demand: a review of empirical research, International Journal of Forecasting, Vol 11, pp 447-475.

2. PODATCI I METODOLOGIJA ISTRAŽIVANJA

S obzirom na dostupnost podataka, kao i postojeću literaturu koja obrađuje tematiku turističke potražnje, za potrebe ovoga rada odabrane su sljedeće varijable: ukupan broj ostvarenih noćenja (domaćih i stranih turista) kao mjera turističke potražnje u RH, bruto domaći proizvod zemalja Europske unije (na bazi 25 članica) kao aproksimacija dohotka turista, indeks potrošačkih cijena (*CPI*) kao aproksimacija troška boravka u turističkoj destinaciji, te tri binarne (*dummy*) varijable kojima je obuhvaćen utjecaj sezone⁵ na turističku potražnju. U analizi su korišteni tromjesečni podatci za razdoblje od I. tromjesečja 1998. do IV. tromjesečja 2007. godine. Nazivi odabranih varijabli, te njihov opis navedeni su u tablici 2.

Tablica 2. Nazivi i opis varijabli uključenih u analizu turističke potražnje

NIGHTS	Ukupan broj noćenja turista u tisućama
GDPEU	BDP zemalja Europske unije na bazi 25 članica, indeks 2005=100
RHCPI	Indeks potrošačkih cijena RH, 2005=100
D1	Binarna varijabla, poprima vrijednost 1 u prvom tromjesečju, u ostalim tromjesečjima poprima vrijednost 0
D2	Binarna varijabla, poprima vrijednost 1 u drugom tromjesečju, u ostalim tromjesečjima poprima vrijednost 0
D3	Binarna varijabla, poprima vrijednost 1 u trećem tromjesečju, u ostalim tromjesečjima poprima vrijednost 0

Izvor: istraživanje autora

Podatci o ukupnom broju noćenja turista preuzeti su iz nekoliko brojeva Mjesečnog statističkog izvješća Državnog zavoda za statistiku, podatci o indeksu potrošačkih cijena preuzeti su iz Priopćenja Državnog zavoda za statistiku (broj 13.1.1/1), a podatci o bruto domaćem proizvodu 25 zemalja članica Europske unije⁶ preuzeti su s web stranice Eurostata.

Analiza je provedena pomoću programskih paketa: *EViews 5.0*, *SAS 9.1* i *Statistica 7*.

3. MODELIRANJE TURISTIČKE POTRAŽNJE RH

3.1 Analiza turističke potražnje metodom višestruke linearne regresije

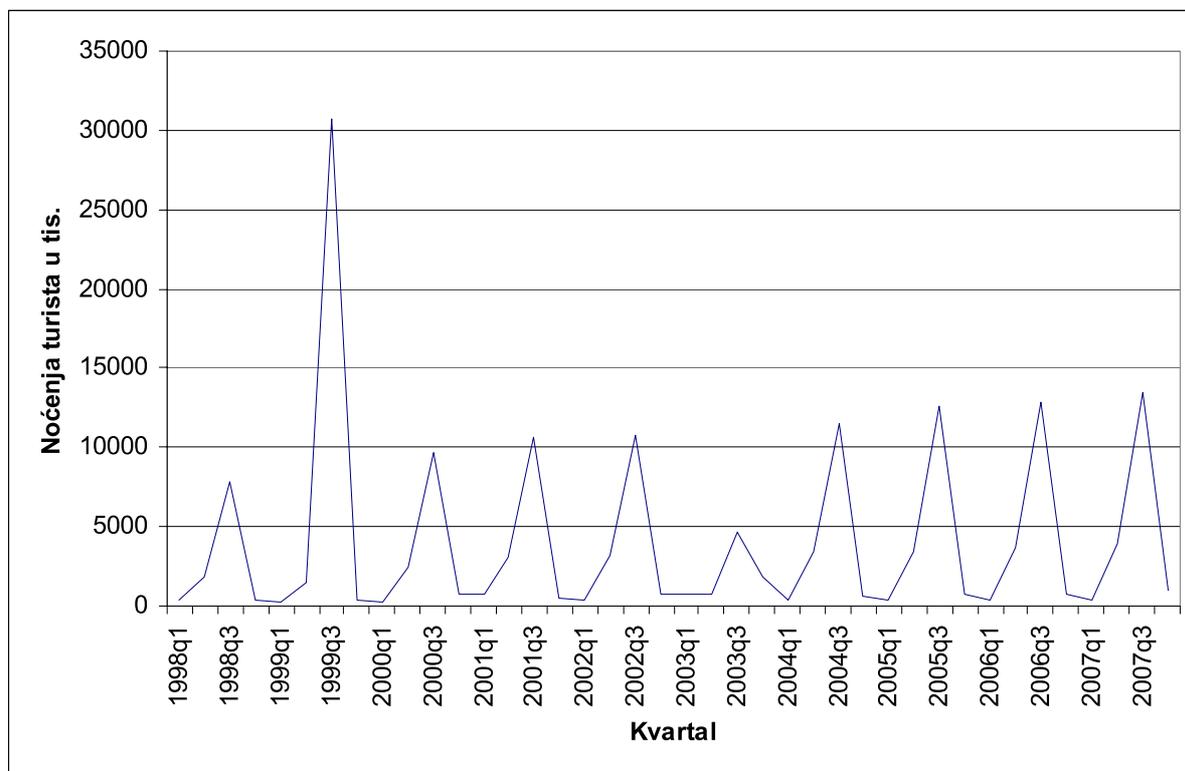
Iako bi analiza modela turističke potražnje zahtijevala sofisticiraniji pristup, poput modela korekcije pogreške (*Error Correction Model*, ECM)⁷, dostupnost relevantnih vremenskih nizova odgovarajuće duljine, ograničuje njegovu upotrebu. U ovome je radu turistička potražnja analizirana na uobičajen način, modelom višestruke linearne regresije. Serija broja noćenja turista prikazana je slikom 1.

⁵ U analizi su korišteni sezonski neprilagođeni podatci, pa je binarnim varijablama (D1, D2, D3) obuhvaćen utjecaj sezone. Kao temeljno tromjesečje odabrano je 4. tromjesečje.

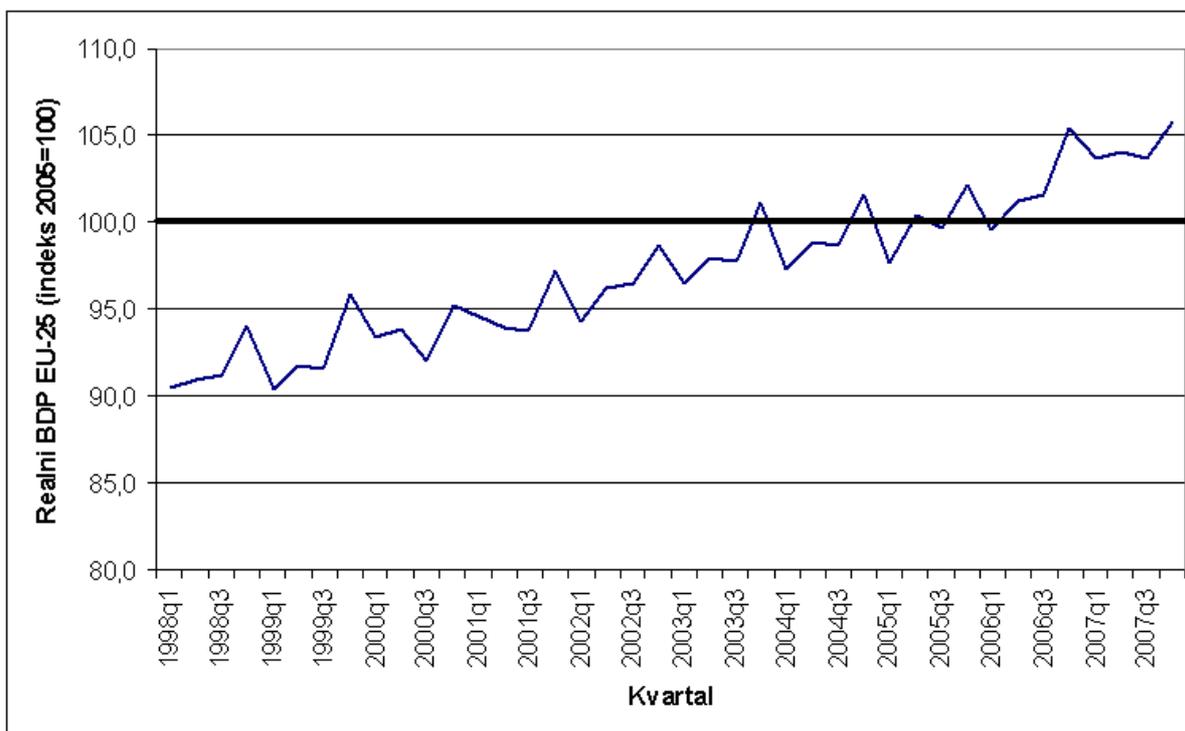
⁶ Indeks bruto domaćeg proizvoda zemalja EU-25 preračunat je sa bazne 2000. na baznu 2005. godinu kako bi bio usporediv s podacima Indeksa potrošačkih cijena u Republici Hrvatskoj.

⁷ Vidjeti Mervar A., Payne J.E. : An Analysis of Foreign Tourism Demand for Croatian Destinations : Long-Run Elasticity estimates, EIZ-WP-0701.

Slika 1. Tromjesečni niz ukupnog broja noćenja turista u RH



Slika 2. Tromjesečni niz indeksa deflaciranog bruto domaćeg proizvoda zemalja članica EU-25



Na osnovi grafičkog prikaza vremenskog niza ukupnog broja noćenja turista, može se uočiti prisutnost sezonalnosti i trenda (vrijednosti niza postupno rastu u promatranom razdoblju). Niža razina ostvarenih noćenja u trećem tromjesečju 2003. godine povezana je sa stagnacijom broja turističkih noćenja koje su, prema podacima Eurostata⁸, zemlje članice EU-25 ostvarile u tom razdoblju. Niz karakterizira i povećan broj ostvarenih noćenja u trećem tromjesečju 1999. godine. Budući da se modelom željelo istražiti kako dohodak turista utječe na potražnju (aproksimiranu ukupnim brojem ostvarenih noćenja), za potrebe modela, indeks bruto domaćeg proizvoda zemalja članica EU-25 na stalnim cijenama 2005. (varijabla GDPEU), deflaciran je indeksom potrošačkih cijena u Republici Hrvatskoj (varijabla RHCPI), te je tako dobivena nova varijabla RGDPEU (realni bruto domaći proizvod zemalja članica EU-25) korištena za daljnju analizu. Varijabla RGDPEU prikazana je slikom 2.

Na osnovi grafičkog prikaza realnog dohotka turista (RGDPEU) može se zaključiti da je vremenski niz sezonskog karaktera, te da varijabla RGDPEU ima izraženi trend, tj. vrijednosti niza rastu tijekom analiziranog perioda.

U svrhu stabiliziranja varijance, provedena je logaritamska transformacija varijabli. Da bi se utvrdio red integriranosti odabranih varijabli, proveden je ADF test (**prošireni Dickey-Fullerov test, Augmented Dickey Fuller test**). ADF test je jedan od testova jediničnog korijena i nastao je kao proširenje *Dickey Fullerova testa* kojim se analizira niz y_t i ispituje njegova stacionarnost. Niz y_t se općenito može predočiti jednadžbom:

$$\Delta Y_t = \alpha_0 + \gamma \cdot Y_{t-1} + \alpha_2 t + \varepsilon_t, \quad (1)$$

pri čemu je ΔY_t diferencirana varijabla Y_t , dok su α_0 , γ , α_2 nepoznati parametri, a $\{\varepsilon_t\}$ je čisti slučajni proces.

Ako je parametar $\gamma = 1 - \rho$ jednak nuli, proces je nestacionaran, odnosno sadrži jedinični korijen $\rho = 1$. Pri tome se razlikuju tri slučaja:

- A. Ako je parametar $\gamma < 0$, proces je stacionaran oko nule (ako je konstantni član α_0 jednak nuli i ako je parametar uz varijablu vrijeme α_2 jednak nuli)
- B. Ako je parametar $\gamma < 0$, te ako je $\alpha_0 \neq 0$, a $\alpha_2 = 0$, proces $\{Y_t\}$ stacionaran je oko konstante.
- C. Ako je $\gamma < 0$, te su $\alpha_0 \neq 0$ i $\alpha_2 \neq 0$, proces $\{Y_t\}$ je trend stacionaran.

Hipoteze testa formuliraju se na sljedeći način:

$$\begin{aligned} H_0: & \rho = 1 \\ H_1: & \rho < 1 \end{aligned} \quad (2)^9$$

tj. nultom se hipotezom pretpostavlja da je proces nestacionaran.

Početno se analizira serija y_t i ispituje njezina stacionarnost. Ukoliko niz y_t nije stacionaran, ispituje se stacionarnost niza prvih diferencija Δy_t itd.

Pođe li se od pretpostavke da je nulta hipoteza o nestacionarnosti procesa istinita, testovne veličine za slučajeve A., B. i C. imaju Dickey-Fullerovu (DF) distribuciju.

⁸ Eurostat, Panorama on Tourism, 2008 Edition

⁹ Hipoteze testa mogu se formulirati i kao : $H_0 \gamma=0$, $H_1 \gamma<0$, pri čemu je $\gamma=\rho-1$.

Ako proces grješaka relacije $\{\varepsilon_t\}$ nije čisti slučajni proces, umjesto Dickey Fullerova testa koristi se njegova modifikacija, ADF test, tako da se u jednadžbu (1) uključuju dodatni članovi kojima se uklanja autokorelacija grješaka relacije. Umjesto da se test provodi polazeći od prvih diferencija niza (tj. od autoregresijskog modela prvog reda, AR(1)), jednadžba (1) se primjenjuje na pomake do uključivo p -tog reda (odnosno polazeći od AR(p) modela)¹⁰:

$$\Delta Y_t = \alpha_0 + \sum_{i=1}^{p-1} \beta_i \Delta Y_{t-i} + \gamma \cdot Y_{t-1} + \alpha_2 t + \varepsilon_t. \quad (3)$$

Rezultati provedenog ADF testa nad logaritmiranim vrijednostima promatranih varijabli predočeni su u tablici 3.

Tablica 3. Rezultati ADF testa za varijable $\ln NIGHTS$ i $\Delta \ln RGDPEU$ (EViews 5.0)¹¹

Null Hypothesis: *lnNIGHTS* has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 2 (Automatic based on SIC, MAXLAG=9)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-10.63114	0.0000
Test critical values: 1% level	-4.226815	
5% level	-3.536601	
10% level	-3.200320	

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

Null Hypothesis: $\Delta \ln RGDPEU$ has a unit root

Exogenous: Constant

Lag Length: 2 (Automatic based on SIC, MAXLAG=9)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-12.71622	0.0000
Test critical values: 1% level	-3.626784	
5% level	-2.945842	
10% level	-2.611531	

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

¹⁰ Usporediti s Bahovec, V., Erjavec, E. (2008): Uvod u ekonometrijsku analizu, Element, Zagreb (u tisku).

¹¹ Null Hypothesis = nulta hipoteza; unit root = jedinični korijen; Exogenous = egzogene, nezavisne varijable; constant = konstanta; Linear trend = linearni trend; Lag length = duljina pomaka; Automatic based on SIC = automatski, baziran na informacijskom kriteriju SIC; MAXLAG = maksimalna duljina pomaka; Augmented Dickey-Fuller test statistic = test veličina proširenog DickeyFullerovog testa; t-Statistics = t- vrijednost; prob = probability = p-vrijednost; Test critical values = kritične granice testa; MacKinnon one-sided p-values = p-vrijednost za jednosmjerni test.

Na osnovi provedenog ADF testa zaključuje se da je varijabla $\ln NIGHTS$ stacionarna. Analizirana varijabla $\ln RGDPEU$ je integrirana prvog reda, no niz prvih diferencija $\Delta \ln RGDPEU$ je stacionaran. Stoga je polazni regresijski model definiran kako slijedi¹²:

$$\ln NIGHTS_t = \beta_0 + \beta_1 \Delta \ln RGDPEU_t + \beta_2 D_2 + \beta_3 D_3 + \varepsilon_t \quad (4)$$

U prvome koraku regresijski je model procijenjen metodom najmanjih kvadrata (OLS), te su provedeni dijagnostički testovi da se ispita jesu li ispunjene polazne pretpostavke o modelu. Provedeni dijagnostički testovi, upućivali su na ispunjenje polaznih pretpostavki.

Budući da se problem promjenjivosti varijance (heteroskedastičnost) često javlja u vremenskim serijama, proveden je i Whiteov test heteroskedastičnosti, koji se ubraja u LM testove (testove *Lagrangeovih multiplikatora*). Prednost je tog testa, u odnosu na ostale LM testove, da se njime ne pretpostavlja normalna distribucija grješaka relacije ε_i , ne pretpostavlja se da su unaprijed poznati regresori z_i u pomoćnoj regresiji, a ne specificira se ni oblik heteroskedastičnosti.

Whiteovim testom se nultom hipotezom pretpostavlja homoskedastičnost, a test se provodi usporedbom LS procjena varijanci u slučaju homoskedastičnosti i heteroskedastičnosti. Uz pretpostavku H_0 o homoskedastičnosti, razlika među dobivenim procjenama bila bi nesigifikantna.

U prvome se koraku metodom najmanjih kvadrata izračunavaju reziduali polaznog regresijskog modela:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (5)$$

U drugome se koraku računa pomoćna regresijska jednadžba u kojoj su $\hat{\varepsilon}_i^2$ vrijednosti zavisne varijable, a regresorske varijable zapravo su regresorske varijable polaznog modela, tj: $x_i, i = 1, 2, \dots, k$ i njihovi kvadrati x_i^2 . Test veličina je nR^2 , pri čemu je n veličina uzorka, a R^2 je koeficijent determinacije pomoćne regresijske jednadžbe. Asimptotska test veličina ima $\chi^2(df = r)$ distribuciju, gdje je r broj nezavisnih varijabli.

U tablici 4¹³ predloženi su rezultati Whiteovog testa za pomoćnu regresiju u kojoj su regresorske varijable regresori polaznog modela i njihovi kvadrati. Na osnovi rezultata provedenog testa zaključuje se da se ne može odbaciti hipoteza H_0 (jer su empirijske ra-

¹² Sezonska *dummy* varijabla D_1 je kao nesigifikantna isključena iz analize.

¹³ White heteroskedasticity test = Whiteov test heteroskedastičnosti; F-statistics = empirijski F-omjer; Prob. F(4,34) = p-vrijednost za teorijski F-omjer; prob. Chi-Square = p-vrijednost za hi-kvadrat distribuciju; test equation = testovna jednadžba; dependent variable = zavisna varijabla; Method: Least squares = metoda najmanjih kvadrata; sample = uzorak; included observations = broj uključenih opažanja; Coefficient = koeficijent; Std. Error = std. pogreška; t-Statistic = t-omjer (t testovna veličina); Prob. = p-vrijednost (empirijska razina sigifikantnosti); R-squared = koeficijent determinacije; Adjusted R-squared = korigirani koeficijent determinacije; S.E. of regression = procijenjena standardna devijacija regresije; Sum squared resid = SSR; Log likelihood = vrijednost logaritamske funkcije najveće vjerodostojnosti (uz pretpostavku o normalnoj distribuiranosti grješaka relacije) izračunata za procijenjene vrijednosti parametara; Durbin-Watson stat = Durbin-Watsonova testovna veličina; Mean dependent var = aritmetička sredina zavisne varijable; S.D. dependent var = standardna devijacija zavisne varijable; Akaike info criterion = Akaikeov informacijski kriterij (AIC); Schwarz criterion = Schwarzov kriterij; F-statistic = F-omjer; Prob(F-statistic) = empirijska razina sigifikantnosti za F-test.

zine signifikantnosti velike u odnosu na uobičajene razine signifikantnosti), tj. varijanca je homoskedastična.

Tablica 4. Rezultati Whiteovog testa heteroskedastičnosti (EViews 5.0)

White Heteroskedasticity test:

F-statistic	0.566805	Prob. F(4,34)	0.688405
Obs*R-squared	2.438060	Prob. Chi-Square(4)	0.655760

Test Equation:

Dependent Variable: RESID²

Method: Least Squares

Sample: 2 40

Included observations: 39

Collinear test regressors dropped from specification

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.109888	0.164273	0.668935	0.5081
$\Delta \ln RGDPEU$	3.166569	2.624854	1.206379	0.2360
$\Delta \ln RGDPEU^2$	46.30020	152.0798	0.304447	0.7626
D2	0.114933	0.180529	0.636644	0.5286
D3	0.099543	0.193990	0.513137	0.6112
R-squared	0.062514	Mean dependent var		0.201998
Adjusted R-squared	-0.047778	S.D. dependent var		0.349420
S.E. of regression	0.357670	Akaike info criterion		0.900795
Sum squared resid	4.349535	Schwarz criterion		1.114072
Log likelihood	-12.56550	F-statistic		0.566805
Durbin-Watson stat	1.008501	Prob(F-statistic)		0.688405

Tablica 5. Rezultati regresijske analize za model (4) dobiveni metodom najmanjih kvadrata (EViews 5.0)¹⁴Dependent Variable: *lnNIGHTS*

Method: Least Squares

Sample (adjusted): 2 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.238709	0.109552	56.94760	0.0000
$\Delta \ln RGDPEU$	8.115044	3.449199	2.352733	0.0244
D2	1.468553	0.187424	7.835443	0.0000
D3	3.111292	0.186733	16.66170	0.0000
R-squared	0.890734	Mean dependent var		7.445419
Adjusted R-squared	0.881368	S.D. dependent var		1.377434
S.E. of regression	0.474429	Akaike info criterion		1.443507
Sum squared resid	7.877917	Schwarz criterion		1.614129
Log likelihood	-24.14839	F-statistic		95.10609
Durbin-Watson stat	2.065196	Prob(F-statistic)		0.000000

Prema navedenom ispisu procijenjena jednadžba glasi:

$$\ln NIGHTS_t = 6.238709 + 8.115044\Delta \ln RGDPEU_t + 1.468553D_2 + 3.111292D_3 \quad (6)$$

(0,109552) (3,449099) (0,187424) (0,186733)

Koeficijent uz prvu regresorsku varijablu je pozitivan i statistički signifikantan na razini značajnosti 5% (empirijska razina značajnosti $p - vrij = 0.0244$). Predznak je u skladu s pretpostavkom i sugerira da je turistička potražnja izrazito osjetljiva na visinu bruto domaćeg proizvoda zemalja Europske unije. Naime, prema ovom modelu, kada se realni dohodak turista (koji je aproksimiran indeksom BDP-a zemalja članica EU-25 deflacioniranim s razinom cijena u Hrvatskoj) poveća za 1%, turistička će potražnja, aproksimirana

¹⁴ Dependent Variable = zavisna varijabla; Method: Least Squares = metoda: najmanjih kvadrata; Sample = uzorak; Included observations = (broj) uključenih opažanja; Coefficient = koeficijent; Std. Error = std. pogreška; t-Statistic = t-omjer (t testovna veličina); Prob. = p-vrijednost (empirijska razina signifikantnosti); R-squared = koeficijent determinacije; Adjusted R-squared = korigirani koeficijent determinacije; S.E. of regression = procijenjena standardna devijacija regresije; Sum squared resid = SSR; Log likelihood = vrijednost logaritamske funkcije najveće vjerodostojnosti (uz pretpostavku o normalnoj distribuiranosti grješaka relacije) izračunata za procijenjene vrijednosti parametara; Durbin-Watson stat = Durbin-Watsonova testovna veličina; Mean dependent var = aritmetička sredina zavisne varijable; S.D. dependent var = standardna devijacija zavisne varijable; Akaike info criterion = Akaikeov informacijski kriterij (AIC); Schwarz criterion = Schwarzov kriterij; F-statistic = F-omjer; Prob(F-statistic) = empirijska razina signifikantnosti za F-test.

ukupnim brojem noćenja u prosjeku porasti za 8,12%. Obje sezonske *dummy* varijable su statistički značajne na bilo kojoj razini signifikantnosti (p -vrijednost=0,0000), a njihov pozitivni predznak ukazuje da će se turistička potražnja zbog utjecaja sezone u prosjeku povećati u predsezoni (travanj, svibanj, lipanj) u odnosu na četvrto tromjesečje za 1,47%, dok će na vrhuncu sezone (srpanj, kolovoz, rujana) turistička potražnja u prosjeku porasti za 3.11%, u odnosu na temeljno četvrto tromjesečje.

Za procijenjeni je model koeficijent determinacije $R^2 = 0.89$, što znači da je modelom protumačeno čak 89% svih odstupanja. Durbin-Watsonov pokazatelj iznosi 2.06. Provedeni jednosmjerni DW-test o pozitivnoj autokorelaciji prvoga reda rezultira prihvaćanjem nulte hipoteze o nepostojanju autokorelacije.

S obzirom da ponekad postoji autokorelacija grješaka relacije reda višeg od 1, promatrana je autokorelacijska funkcija reziduala i proveden Ljung-Box Q-test o autokorelaciji višega reda.

U statističkom softveru EViews dan je prikaz autokorelacijske (AC) i parcijalne autokorelacijske funkcije (PAC), te su izračunate Ljung-Boxove Q-testovne veličine i njihove empirijske razine signifikantnosti (Prob), kao što je to prikazano na slici 3.

Slika 3. Autokorelacijska i parcijalna autokorelacijska funkcija reziduala, Q-test veličine i empirijske razine signifikantnosti

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.051	-0.051	0.1099	0.740
		2	-0.134	-0.137	0.8883	0.641
		3	0.037	0.022	0.9477	0.814
		4	-0.009	-0.025	0.9516	0.917
		5	0.004	0.011	0.9526	0.966
		6	0.211	0.211	3.1059	0.795
		7	0.089	0.123	3.4996	0.835
		8	-0.066	0.005	3.7249	0.881
		9	-0.219	-0.225	6.2927	0.710
		10	-0.014	-0.071	6.3033	0.789
		11	0.056	-0.012	6.4806	0.839
		12	-0.058	-0.103	6.6774	0.878
		13	0.062	0.026	6.9141	0.907
		14	0.330	0.395	13.873	0.459
		15	-0.199	-0.005	16.498	0.350
		16	-0.184	-0.117	18.848	0.277

Hipoteze za Ljung-Boxov Q- test su sljedeće:

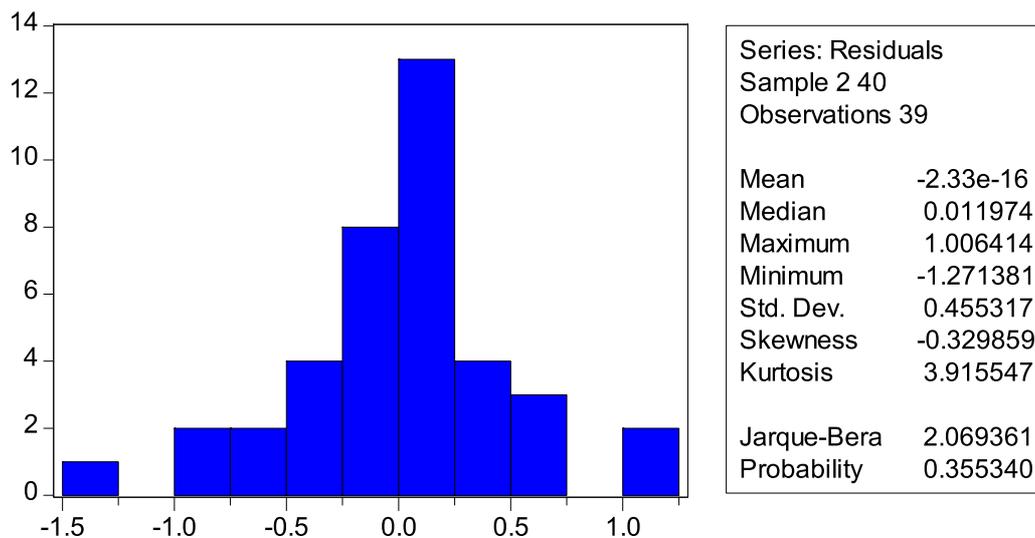
$$H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0 \quad (7)$$

$$H_1 : \exists \rho_j \neq 0, j = 1, 2, \dots, k$$

Sadržaj nulte hipoteze je da nema autokorelacije do reda k (tj. da su svi koeficijenti autokorelacije jednaki nuli, te da je niz grješaka relacije $\{\varepsilon_t\}$ čisti slučajni proces ili bijeli šum). S obzirom da su za svaki pomak k , Q-testovne veličine nesignifikantne, zaključuje se da ne postoji problem autokorelacije grješaka relacije do reda k .

Također je ispitana pretpostavka o normalnosti grješaka relacije Jarque-Bera testom, koji koristi koeficijent asimetrije i koeficijent zaobljenosti reziduala procijenjenih metodom najmanjih kvadrata. Pri tome se ispituje odstupaju li procijenjene veličine značajno od vrijednosti tih mjera za normalnu distribuciju. Kao što se vidi na slici 4, za procijenjeni model testovna je veličina $JB=2,069361$, a pridružena empirijska razina signifikantnosti (Probability) iznosi 0.355340, pa se uz bilo koju uobičajenu razinu signifikantnosti ne može odbaciti nulta hipoteza o normalnoj distribuiranosti grješaka relacije.

Slika 4. Histogram rezidualnih odstupanja i rezultati Jarque-Bera testa¹⁵



Nadalje, ispitana je pretpostavka o približnoj linearnoj zavisnosti regresorskih varijabli (multikolinearnosti). Najjednostavniji način utvrđivanja ozbiljnog problema multikolinearnosti je ispitivanje standardnih pogriješaka procijenjenih parametara.

Varijanca od $\hat{\beta}_j$ dana je izrazom:

$$\text{var}(\hat{\beta}_j) = \sigma^2 (X'X)^{-1}_{jj} = \frac{\hat{\sigma}^2}{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \cdot (1 - R_j^2)}, \quad (8)$$

pri čemu je R_j^2 koeficijent determinacije u modelu multiple linearne regresije u kojem je j -ta regresorska varijabla zavisna, a preostali regresori nezavisne varijable. Ako postoji uska linearna povezanost između varijable x_j i preostalih regresorskih varijabli, koeficijent determinacije je velik (blizu 1), pa je i varijanca od $\hat{\beta}_j$ velika. Među standardnim pokazateljima multikolinearnosti u programskim paketima su faktor inflacije varijance, VIF (*Variance Inflation Factor*) ili ekvivalentni pokazatelj TOL (Tolerance), definirani izrazom:

¹⁵ Mean = aritmetička sredina; Median = medijan; maximum = najveća vrijednost; Minimum = najmanja vrijednost, Std. Dev. = standardna devijacija; Skewness = koeficijent asimetrije; Kurtosis = zaobljenost; Jarque-Bera = test- veličina; probability = p-vrijednost.

$$VIF_j = \frac{1}{1 - R_j^2} \quad j = 1, 2, \dots, p \quad TOL_j = \frac{1}{VIF_j} = 1 - R_j^2 . \quad (9)$$

Ozbiljan problem multikolinearnosti je prisutan ako je $R_j^2 > 0.8$, odnosno $VIF_j > 5$, ili ekvivalentno $TOL_j < 0.2$.

Naziv "faktor inflacije varijance" proizlazi iz činjenice da u slučaju visoke korelacije regresorske varijable X_j s ostalim regresorskim varijablama, koja rezultira koeficijentom determinacije $R_j^2 \approx 1$, dolazi do povećanja ("inflacije") varijance od $\hat{\beta}_j$. Pokazatelji VIF i TOL predočeni u tablici 5. upućuju na zaključak da u modelu ne postoji problem multikolinearnosti.

Tablica 6. Pokazatelji multikolinearnosti

Varijabla	TOL	VIF
ln <i>NIGHTS</i>	0.94619	1.05687
D2	0.865363	1.155584
D3	0.962481	3.116945

Izvor: izračun autora

Osim izračunavanja navedenih standardnih pokazatelja multikolinearnosti (VIF i TOL), analiza problema multikolinearnosti zasniva se i na uporabi svojstvenih vrijednosti matrice $(X'X)^{-1}$, pomoću kojih se određuje tzv. kondicioni broj, odnosno kondicioni indeks¹⁶.

Kondicioni broj (CN) i kondicioni indeks (CI) matrice $(X'X)^{-1}$ ubrajaju se među pokazatelje multikolinearnosti, jer su to mjere osjetljivosti procjena u regresiji na male promjene u podacima. Kondicioni broj definiran je kao drugi korijen omjera najveće i najmanje svojstvene vrijednosti, tj.:

$$CN = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} , \quad (10)$$

a kondicioni indeks je određen kao drugi korijen omjera najveće svojstvene vrijednosti i odabrane i -te svojstvene vrijednosti λ_i :

$$CI_i = \sqrt{\frac{\lambda_{\max}}{\lambda_i}} , \quad i = 1, 2, \dots, k . \quad (11)$$

¹⁶ Skalar λ_i i vektor v_i za koje vrijedi $(X'X)^{-1}v_i = \lambda_i v_i$ zovu se svojstvena vrijednost i svojstveni vektor matrice $(X'X)^{-1}$.

Poželjno je da kondicioni broj bude što manji, tj. što bliži 1. Smatra se da ne postoji problem multikolinearnosti ako su kondicioni indeksi (CI) manji od 30. Izračunavanje svojstvenih vrijednosti provodi se metodom glavnih komponenata. Matrica $(X'X)^{-1}$ se u prvom koraku transformira tako da ima jedinice na glavnoj dijagonali, odnosno tako da je zbroj dijagonalnih elemenata jednak $k+1$ (tj. trag matrice $tr(X'X)^{-1} = k+1$).

U drugom se koraku singularnom dekompozicijom matrica $(X'X)^{-1}$ svodi na dijagonalnu matricu.¹⁷

U programskom paketu SAS 9.1 u ispisu za utvrđivanje multikolinearnosti predočene su svojstvene vrijednosti glavnih komponenata, kondicioni indeksi te proporcija varijance svake ulazne varijable protumačena pojedinom glavnim komponentom. Kao što se može uočiti iz ispisa, kondicioni indeksi su mali te ne postoji problem multikolinearnosti.

Tablica 7. Pokazatelji multikolinearnosti CI (ispis SAS 9.1)

Collinearity Diagnostics						
Condition -----Proportion of Variation-----						
Number	Eigenvalue	Index	Intercept	dlgdp	D2	D3
1	1.78287	1.00000	0.12382	0.04358	0.10240	0.06205
2	1.18749	1.22531	0.01126	0.29162	0.08288	0.23985
3	0.74694	1.54496	0.00484	0.65855	0.27884	0.11355
4	0.28270	2.51127	0.86008	0.00626	0.53588	0.58456

Izvor: izračun autora

Budući da su svi statistički pokazatelji adekvatnosti izbora modela zadovoljavajući, model (6) odabire se kao konačni model kojim je moguće opisati turističku potražnju u Republici Hrvatskoj aproksimiranu ukupnim brojem ostvarenih noćenja.

¹⁷ Za kvadratnu matricu $(X'X)^{-1}$ postoje matrice V i V' tako da vrijedi: $(X'X)^{-1} = V\Lambda^2V'$. Pritom je V ortogonalna matrica ($VV' = V'V = I$) čiji su stupci svojstveni vektori matrice $(X'X)^{-1}$, a Λ^2 je dijagonalna matrica svojstvenih vrijednosti matrice $(X'X)^{-1}$. Navedena dekompozicija matrice $(X'X)^{-1}$ zove se singularna dekompozicija. Svojstvene vrijednosti λ_i izražavaju mjeru disperzije u prostoru u kojem su koordinatne osi glavne komponente. To su međusobno ortogonalne (nekorelirane) varijable, koje reprezentiraju smjer maksimalnog varijabiliteta u prostoru X . Prva je glavna komponenta s najvećom varijancom, druga glavna komponenta ortogonalna je na prvu i ima najveću varijancu iza prve komponente itd. Proporcija ukupne varijance protumačena i -tom glavnim komponentom jednaka je $\lambda_i/(k+1)$.

4. ZAKLJUČAK

Cilj ovoga rada bio je procijeniti ekonometrijski model turističke potražnje u Republici Hrvatskoj i prepoznati koje su determinante u njoj podlozi od najvećeg utjecaja.

Rezultati analize su pokazali da realni dohodak turista kao i sezonski faktori (i to treće tromjesečje) u najvećoj mjeri utječu na potražnju u turizmu. Navedeni rezultati su u skladu s recentnom literaturom iz ovoga područja.

Prema procijenjenom modelu, turistička potražnja u Republici Hrvatskoj izrazito je osjetljiva na bruto domaći proizvod zemlja Europske unije, pa bi nositelji ekonomske politike trebali biti svjesni potencijalnog učinka što ga promjene ekonomske aktivnosti u zemljama članicama mogu imati na sektor turizma, kao i na cjelokupnu ekonomsku aktivnost u Hrvatskoj.

Rezultati analize sugeriraju i izrazitu osjetljivost turističke potražnje na sezonu. Pokazalo se da su binarne varijable za drugo i treće tromjesečje u cijelom promatranom razdoblju statistički značajne na bilo kojoj razini signifikantnosti, a najveći utjecaj na porast turističke potražnje ima treće tromjesečje. Procijenjeni model je podvrgnut odgovarajućim dijagnostičkim testovima i pokazao se zadovoljavajućim.

Kao restrikcije analiziranog modela, mogu se navesti kratkoća vremenske serije, pa modelu valja pristupiti kao inicijalnom pokušaju za razumijevanje determinanti u podlozi turističke potražnje u Hrvatskoj. Daljnje analize, s duljim serijama, kao i usporedbe turističke potražnje u regiji trebale bi se provesti u nastavku istraživanja.

LITERATURA

1. Aczel, A. D., Souderpandian, J.: "Complete Business Statistics", 6th Edition, McGraw-Hill, New York, 2006.
2. Bahovec, V., Erjavec, N.: Uvod u ekonometrijsku analizu, Element, Zagreb (u tisku).
3. Enders, W.: "Applied Econometric Time Series", 2nd Edition, John Wiley & Sons, New York, 2004.
4. Eurostat, "Panorama on Tourism", 2008 Edition.
5. Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L.: "Multivariate Data Analysis", 6th Edition, Pearson, Prentice Hall, Upper Saddle River, NJ, 2006.
6. Johnson, R. A., Wichern, D., W.: Applied Multivariate Statistical Analysis, 6th Edition, Pearson, Prentice Hall, Upper Saddle River, NJ, 2007.
7. Manly, B. F. J.: "Multivariate Statistical Methods", Chapman & Hall/CRS, London, 2005.
8. Mervar, A., Payne, J.E.: An Analysis of Foreign Tourism Demand for Croatian Destinations: Long-Run Elasticity Estimates, Ekonomski institut Zagreb, Working papers, EIZ-WP-0701.
9. Payne, J. E., Mervar, A.: A note on modelling tourism revenues in Croatia, "Tourism Economics", 2002, 8 (1), 103-109.

10. Preez, J., Witt, S. F.: Univariate versus multivariate time series forecasting: an application to international tourism demand, "International Journal of Forecasting", 19 (2003), 435-451.
11. Stučka, T.: A comparison of two econometrics models (OLS and SUR) for forecasting Croatian tourism arrivals, "Working Papers", Croatian National Bank, July 2002.
12. Turner, L. W., Witt, S. F.: Forecasting tourism using univariate and multivariate structural time series models, "Tourism Economics", 2001, 7 (2), 135-147.
13. Witt, S. F. and Witt, C. A., Forecasting tourism demand: a review of empirical research, "International Journal of Forecasting", Vol. 11, pp. 447-475.
14. http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1090,30070682,1090_33076576&dad=portal&schema=PORTAL.eurostat
15. <http://www.dzs.hr>